The US Army's Center for Strategy and Force Evaluation

MEMORANDUM REPORT
CAA-MR-97-3

THE ADVANTAGE PARAMETER: A COMPILATION OF PHALANX ARTICLES DEALING WITH THE MOTIVATION AND EMPIRICAL DATA SUPPORTING USE OF THE ADVANTAGE PARAMETER AS A GENERAL MEASURE OF COMBAT POWER

JULY 1997

CONCEPTS ANALYSIS AGENCY
UNITED STATES ARMY

PREPARED BY
TACTICAL ANALYSIS DIVISION

US ARMY CONCEPTS ANALYSIS AGENCY
8120 WOODMONT AVENUE
BETHESDA, MARYLAND 20814-2797

CONCEPTS ANALYSIS AGENCY

DISCLAIMER

The findings of this report are not to be construed as an official Department of the Army position, policy, or decision unless so designated by other official documentation. Comments or suggestions should be addressed to:

Director
US Army Concepts Analysis Agency
ATTN: CSCA-TA
8120 Woodmont Avenue
Bethesda, MD 20814-2797

CAA Memorandum Reports are used for the convenience of the sponsors or the analysts to record substantive work done in quick reaction studies and major interactive technical support activities; to make available preliminary and tentative results of analyses or of working group and panel activities; or to make a record of conferences, meetings, or briefings, or of data developed in the course of an investigation. Memorandum Reports are reviewed to assure that they meet high standards of thoroughness, objectivity, and sound analytical methodology.

CAA-MR-97-3

REPORT DOCUMENTATION PAGE	*Form Approved* *OPM NO. 0704-0188*

Public reporting burden for this collection information is estimated to 1 hour per response, including the time for reviewing instructions, searching existing data sources gathering and maintaining the data needed, and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information. Including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of information and Regulatory Affairs, Office of Management and Budget, Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1997	Final, May - July 1997

4. TITLE AND SUBTITLE	5. FUNDING NUMBER
The Advantage Parameter: A Compilation of Phalanx Articles ... Supporting Use of the Advantage Parameter as a General Measure of Combat Power	
6. AUTHOR(S) Dr. Robert L. Helmbold	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **US Army Concepts Analysis Agency** **8120 Woodmont Avenue** **Bethesda, MD 20814-2797**	8. PERFORMING ORGANIZATION REPORT NUMBER CAA-MR-97-3
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/ MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILBILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for public release; dissemination unlimited	A

13. ABSTRACT (Maximum 200 words)

This paper is a compilation of previously published articles on the theoretical motivation and empirical data supporting use of the advantage parameter as a general measure of combat power. Some additional material has been added to provide suitable links and references to some of the material on which these articles are based.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Victory, combat power, combat effectiveness, winning, combat superiority, measures of effectiveness, figures of merit, personnel, attrition, combat, battle, casualties, losses, trends, history	73
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500
Standard Form 298

MEMORANDUM REPORT
CAA-MR-97-3

THE ADVANTAGE PARAMETER: A COMPILATION OF PHALANX ARTICLES DEALING WITH THE MOTIVATION AND EMPIRICAL DATA SUPPORTING USE OF THE ADVANTAGE PARAMETER AS A GENERAL MEASURE OF COMBAT POWER

July 1997

Prepared by

TACTICAL ANALYSIS DIVISION

US Army Concepts Analysis Agency
8120 Woodmont Avenue
Bethesda, Maryland 20814-2797

DEPARTMENT OF THE ARMY
US ARMY CONCEPTS ANALYSIS AGENCY
8120 WOODMONT AVENUE
BETHESDA, MARYLAND 20814-2797

REPLY TO
ATTENTION OF:

CSCA-TA (5-5d)

23 JUN 1997

MEMORANDUM FOR Deputy Under Secretary of the Army, (OR), Headquarters, Department of the Army, Washington, DC 20310

SUBJECT: The Advantage Parameter: A Compilation of Phalanx Articles Dealing with the Motivation and Empirical Data Supporting Use of the Advantage Parameter as a General Measure of Combat Power

1. The U.S Army Concepts Analysis Agency (CAA) is pleased to publish this memorandum report by Dr. Robert L. Helmbold. It provides a useful summary of the theoretical motivation and empirical data that support use of the advantage parameter as a general measure of a force's overall combat power or military effectiveness. This overall measure of effectiveness has proven useful in a number of CAA analyses. We believe its wider use would be beneficial and that it could be used to replace, or at least supplement, many of the various measures of combat effectiveness currently being used within the Army. Properly used, such measures of overall combat power could improve and standardize the interpretation of a broad spectrum of U.S. Army models, wargames, studies, and analyses. Accordingly, wide dissemination of this work is encouraged.

2. Questions or inquiries should be directed to the Tactical Analysis Division, U.S. Army Concepts Analysis Agency (CSCA-TA), 8120 Woodmont Avenue, Bethesda, MD 20814-2797, (301) 295-1611 or DSN 295-1611.

E. B. Vandiver Jr

E. B. VANDIVER III
Director

PREFACE

Over the period December 1992 to January 1997, 10 articles on the advantage parameter either appeared in or were submitted for publication to the author's Analyst's Corner column of PHALANX, the Military Operations Research Society's newsletter. Considered collectively, these articles provide a summary of the motivation and empirical data supporting use of the advantage parameter as a general measure of the combat effectiveness or power of a military force. The impact of these articles was somewhat diluted because they appeared individually and sporadically over an extended period of time. Here we have collected these articles into a single document that makes it easier to understand the interrelationships among them, provides a useful unified reference to them, and affords a general overview or summary of the current state of research on the advantage parameter.

The articles can conveniently be divided into two main parts—theoretical motivation and empirical justification. The part dealing with theoretical motivation has very limited objectives. It is used only to suggest a particular formulation for an advantage parameter that potentially could have practical application. However, no theoretical considerations could possibly determine whether or not a theoretically-motivated advantage parameter realizes its potential in actual practice. The only way to determine that is to appeal to empirical evidence from historical cases. Nevertheless, hardly any of the usual theoretical measures of combat effectiveness have been tested empirically. Accordingly, the second (and more important) part on empirical justification analyzes several different and largely independent data bases, and convincingly demonstrates that each of them strongly supports the practical use of the advantage parameter as a general measure of combat power or military effectiveness.

The articles are reproduced here essentially as they originally appeared. However, we have taken advantage of this opportunity to clarify and standardize the mathematical notation, as well as to format the material to adhere more closely to CAA's publication guidelines. We have also added some material to provide a few links and connections to the most important works from which the PHALANX articles were derived. The interested reader can use these to explore more deeply into the theoretical and empirical basis supporting use of the advantage parameter in situations of practical military affairs. We have also added some introductory material on kernel smoothing, a statistical technique used here as a descriptive statistical method for presenting some of the empirical data, as an aid to readers unfamiliar with that technique.

The present report is provided in the spirit of the following remarks of John I. Alger, which appeared in *The Quest for Victory: The History of the Principles of War*, Greenwood Press, Westport, Connecticut, 1982, pg 173: "The modern battlefield is a confused chorus of cacaphony. Filthy sweat, painful exhaustion, utter misery, sickness, and death generally characterize the experiences upon that field. For the individual participant, survival is often goal enough, but for the professional leader and the directors of war, the goal must be victory over the trials of the operations, the confusion of the battle, and ultimately over the enemy forces. The task is formidable, and throughout the written history of war, hardly a leader exists who has not known the bitterness of defeat. But in spite of the mercurial nature of the fortunes of war and of the knowledge that for every victor there must be a vanquished, writers and teachers, military and civilian, have sought to identify the elements of victory and to ensure ordered thinking where physical confusion abounds. Their efforts often resulted in the charge of pedantry, a scathing criticism from the profession of practical men of action; but the cost of defeat is too great to rebuke any effort that might contribute to success in battle."

THE ADVANTAGE PARAMETER: A COMPILATION OF PHALANX ARTICLES DEALING WITH THE MOTIVATION AND EMPIRICAL DATA SUPPORTING USE OF THE ADVANTAGE PARAMETER AS A GENERAL MEASURE OF COMBAT POWER

THE REASON FOR PERFORMING THIS RESEARCH was that these articles provide a summary of the motivation and empirical data supporting use of the advantage parameter as a general measure of the combat effectiveness or power of a military force. Publishing these articles as a collection makes it easier to understand the interrelationships among them, provides a useful unified reference to them, and affords a general overview or summary of the current state of research on the advantage parameter.

THE SPONSOR was the Director, US Army Concepts Analysis Agency.

THE OBJECTIVE was to provide in a single document a summary of the motivation and empirical data supporting use of the advantage parameter as a general measure of combat effectiveness.

THE SCOPE OF THE RESEARCH involves republishing the articles published in or submitted to the PHALANX from December 1992 through January 1997, with some relatively minor emendations and additions.

THE BASIC APPROACH used in this research is to formulate an advantage parameter that appears on theoretical grounds to have good potential as a practical measure of combat effectiveness, and then to subject it to rigorous testing against the empirical data currently available at CAA.

THE PRINCIPAL FINDINGS of the work reported herein are that:

a. The proposed advantage parameter is well-motivated on theoretical grounds. However, these alone are not sufficient to justify its use in practical military applications.

b. Practical use of the proposed advantage parameter is strongly supported by all of the empirical data currently available to CAA.

THE RESEARCH EFFORT was directed by Dr. Robert L. Helmbold, Tactical Analysis Division.

COMMENTS AND SUGGESTIONS may be sent to the Director, US Army Concepts Analysis Agency, ATTN: CSCA-TA, 8120 Woodmont Avenue, Bethesda, Maryland, 20814-2797.

CONTENTS

TABLES

TABLE **Page**

FIGURES

FIGURE

CHAPTER 1

EXECUTIVE SUMMARY

1-1. BACKGROUND

a. Over the period December 1992 to January 1997, 10 articles on the advantage parameter appeared in or were submitted for publication to the author's Analyst's Corner column of PHALANX, the Military Operations Research Society's newsletter. Considered collectively, these articles provide a summary of the motivation and empirical data supporting use of the advantage parameter as a general measure of the combat effectiveness or power of a military force. The impact of these articles was somewhat diluted because they appeared individually and sporadically over an extended period of time. Here we have collected these articles into a single document that makes it easier to understand the interrelationships among them, provides a useful unified reference to them, and affords a general overview or summary of the current state of research on the advantage parameter.

b. The articles can conveniently be divided into two main parts—theoretical motivation and empirical justification. The part dealing with theoretical motivation has very limited objectives. It is used only to suggest a particular formulation for an advantage parameter that potentially could have practical application. However, no theoretical considerations could possibly determine whether or not a theoretically-motivated advantage parameter realizes its potential in actual practice. The only way to determine that is to appeal to empirical evidence from historical cases. Nevertheless, hardly any of the usual theoretical measures of combat effectiveness have been tested empirically. Accordingly, the second (and more important) part on empirical justification analyzes several different and largely independent data bases, and convincingly demonstrates that each of them strongly supports the practical use of the advantage parameter as a general measure of combat power or military effectiveness.

1-2. OBJECTIVE. The main objective of this report is to provide in a single document a summary of the motivation and empirical data supporting use of the advantage parameter as a general measure of combat effectiveness.

1-3. SCOPE. Each of the 10 articles that appeared in or were submitted to PHALANX for publication in the Analyst's Corner column is reproduced here essentially as it originally appeared. However, we have taken advantage of this opportunity to clarify and standardize the mathematical notation, as well as to format the material to adhere more closely to CAA's publication guidelines. We have also added some material to provide a few links and connections to the most important works from which the PHALANX articles were derived. The interested reader can use these to explore more deeply into the theoretical and empirical basis supporting use of the advantage parameter in situations of practical military affairs. We have also added (in Appendix D) some introductory material on kernel smoothing, used here as a descriptive statistical method for presenting some of the empirical data, for the convenience of readers unfamiliar with that technique. Some background material on the development of the

advantage parameter concept and its validation against historical data is also provided in Appendix D.

1-4. FINDINGS, OBSERVATIONS, AND CONCLUSIONS. The principal findings of the work reported herein are that:

a. The proposed advantage parameter is well-motivated on theoretical grounds. However, these alone are not sufficient to justify its use in practical military applications.

b. Practical use of the proposed advantage parameter is strongly supported by all of the empirical data currently available to CAA.

c. Earlier works that included some effort on the advantage parameter (see Helmbold [1986] and Helmbold [1990]), concluded that:

(1) The advantage parameter should be used as a payoff function. Each side can increase its relative advantage parameter only at the expense of decreasing its opponent's advantage parameter by the same amount. Thus, each side seems to be in a zero-sum game with the advantage parameter as the payoff function that each is striving to optimize. Accordingly, the advantage parameter should be used in studies and analyses as the payoff function or figure or merit for assessing the value of alternative organizations, tactics, equipment, and force mixes. Soldiers and commanders should be taught in their service schools, academies, war colleges, and staff colleges that high values of the advantage parameter are strongly associated with winning battles—and therefore that increasing, or at least appraising, the value of their advantage parameter could be very important in battles and similar tactical engagements. Perhaps computation of the advantage parameter during the early stages of a battle would improve tactical decisions for the conduct of the rest of the battle. If, at an early stage in the battle, the advantage parameter is found to be unfavorable, then the commander should either immediately seek additional support or other means for improving his advantage parameter, or else he should attempt to break off the current engagement as expeditiously as possible and to find more favorable circumstances for engaging the enemy.

(2) The relation of the advantage parameter to victory in battle can be used for historical criticism and analysis. For example, if a force that had a large probability of winning the battle reportedly lost it, that is sufficient reason to review the evidence more closely to determine whether the historical reports are accurate and, if they are, what caused this unexpected and unusual turn of events. The advantage parameter may also be used to rate the performance of historical captains—commanders that were consistently able to achieve favorable advantage parameter values would rate highly. A similar rating system for friendly units in time of war may be possible, provided, of course, accurate and dependable data on friendly and enemy strengths and losses are available.

(3) If computed at frequent times during the course of the action, the relation of the advantage parameter to victory in battle could be used as an index of whether the battle currently is going in favor of one side or the other. A specific application of this idea to simulating a theater commander's decision to escalate to weapons of mass destruction (or to de-escalate from them) has been proposed for use in war games (see CAA [1985]).

(4) The relation of the advantage parameter to victory in battle can also be used to test wargames and theories of combat for realism. If the wargame or theory is inconsistent with the empirically observed relationship of the advantage parameter to victory, then that wargame or theory of combat is highly suspect and its results should be used with extreme caution.

(5) We also note that Helmbold [1990] has found that the rate of advance in battles is strongly influenced by the value of the advantage parameter. This observation is currently used by CAA to simulate the rate of advance in its wargames.

CHAPTER 2

THEORETICAL MOTIVATION

2-1. INTRODUCTION. This chapter presents the four articles (numbers 1 through 4) dealing with the theoretical motivation for choosing a particular formulation for the advantage parameter.

2-2. ARTICLE NUMBER 1

(This article was published in the December 1992 issue of PHALANX.)

This is the first column of a new Phalanx Department. The Combat Analysis Department will undertake to explore areas related to the theory of combat operations. The following indicates the kinds of articles solicited for publication in this department.

1. Mathematical models of combat, such as Lanchester's and related deterministic or probabilistic models. The emphasis in this area is on relatively simple models, in the sense that they use aggregate measures of force strength and involve only a few parameters.

2. Statistical treatments of quantitative and other data from historical conflict situations, or from conflict games or simulations that are substantially more complex than the simple models mentioned in item 1.

3. Basic issues involving areas 1 and 2. This includes comparisons and contrasts of these areas with each other. It also includes discussion of the comparative merits of alternative methods of analysis. Observations and essays on the philosophical and scientific underpinnings of combat models and theory, to include the exposition of common errors or misinterpretations, are welcome. Discussions of the relation of combat models and wargames to wider, but closely allied, fields of scientific and military analysis (such as the principles of war, military tactics and strategy, force structure analysis, sociology, economics, international relations and political science, *etc.*) are also welcome, provided they are pertinent to areas 1 and 2, above.

To get things rolling, I will begin with an observation on the relationship (or non-relationship) of Lanchester's square law differential equations to real battles. Several derivations of these equations have been put forward. Many of them are mentioned in Prof. James G. Taylor's thorough two-volume monograph on Lanchester models of warfare Taylor [1983, Vol I, pp 159-166]. One such derivation, which we adapt from Taylor [1983] (and which he, in turn, adapted from Herbert K. Weiss) uses the following postulates:

1. Two forces are engaged in a fire fight, and they are homogeneous forces. That is, each force consists of but one type of fire unit.

2. Each fire unit on either side is within range of all fire units on the other side.

3. The effects of successive rounds are independent.

4. Each fire unit is sufficiently aware of the locations and conditions of all enemy fire units that it concentrates its fire on only the surviving ones. It does this by selecting one of them, which it neutralizes after a fixed time that does not depend on the other events that occur during the engagement. Once that target is neutralized, the fire unit then searches for another target, which it finds after a fixed time that is independent of the other events that occur during the engagement.

5. Each side uniformly distributes its fire over all of the surviving enemy fire units.

From these postulates, Lanchester's square law differential equations,

$$\left.\begin{aligned} x' &= -Dy \\ y' &= -Ax \end{aligned}\right\}$$

can be derived, where x and y are the surviving strengths of the two sides as of time t into the battle, primes denote differentiation with respect to time, and A and D are constant attrition parameters (that is, during the battle they do not depend upon the time t nor on the force strengths x and y).

Now, it is clear that real combat actions seldom (if ever) satisfy the stated postulates. It is tempting to conclude from this fact that Lanchester's square law attrition equations are not valid for real combat actions. However, this argument cannot be sustained. The reason why it cannot is that it depends on the stated postulates being logically *necessary* conditions for the validity of Lanchester's square law. But, while it is true that the stated postulates have been shown to be logically *sufficient* for the validity of Lanchester's square law, no one has claimed that they are logically *necessary* for its validity. So the reasoning is flawed. That is, although the stated conclusion may be a true statement, its truth cannot logically be deduced from the fact that the stated postulates are not valid for real combat actions.

To correct the argument, we would first have to establish logically *necessary* conditions for the validity of Lanchester's square law attrition equations, and then show that at least one of these necessary conditions is not valid for real combat actions. I leave the first step of this argument as a problem to be addressed in the next column—what are the *necessary and sufficient* conditions for the validity of Lanchester's square law attrition equations?

2-3. ARTICLE NUMBER 2

(This article was published in the March 1993 issue of PHALANX.)

In our last column, we pointed out that some derivations of Lanchester's square law attrition equations provide sufficient conditions for its validity. We also observed that some analysts have mistaken these for necessary conditions and hastily concluded that these equations could not possibly be valid in the real world. The question was then raised—What are the necessary and sufficient conditions for the validity of Lanchester's square law attrition equations?

Our answer to this question is as follows. The necessary and sufficient condition for the validity of Lanchester's square law is that each side's attrition rate (*i.e.*, losses per unit time) be proportional to the other side's strength.

Sufficiency follows by observing that a direct translation of the words into mathematical symbols gives the square law attrition equations:

$$\left.\begin{aligned} x' &= -Dy \\ y' &= -Ax \end{aligned}\right\}$$

where A and D are the proportionality factors, x and y are the surviving strengths of the two sides as of time t into the battle, and primes denote differentiation with respect to time. The attrition coefficients A and D are constants independent of the time t and of the force strengths x and y.

Necessity is shown by observing that these equations express the notion that each side's attrition rate is proportional to the other side's strength.

The above argument does not depend upon the detailed actions of the opposed forces. That is both a strength and a weakness. In many investigations, it is useful to avoid a formulation that depends on the specific details of the attrition process. The argument's weakness is that it does not inform us which detailed attrition processes will or will not lead to a strict (or at least approximate) proportionality of each side's losses to the opposing side's strength.

How can we decide whether or not the square law attrition equations are valid in real combat operations? Well, one approach is to consider the empirical evidence for or against them. In order to do that, we must find out whether the predicted attrition values agree with the observed ones. An even better approach is to consider several alternative attrition laws (such as the square law, Lanchester's linear law, a mixed law, and so forth) to see which best fits the empirical data. This raises an important question: What sort of data do we need in order to test the fit of simple attrition laws? Evidently this will differ depending on the attrition law under consideration.

The examples given below are provided solely to illustrate a few types of simple combat attrition law. In their simplest form, these laws or equations deal only with "homogeneous forces," *i.e.*, the forces are treated as consisting of only a single type of element. In that case, the equations usually are special cases of the following:

$$\left.\begin{aligned} dx/dt &= -DF(x,y) + R_x \\ dy/dt &= -AG(x,y) + R_y \end{aligned}\right\}$$

with the initial conditions:

$$x(0) = x_0 \text{ and } y(0) = y_0.$$

Here x and y are the surviving strengths of the attacker and defender at time t after the start of the battle, x_0 and y_0 are their respective strengths at the start of the battle, A and D are their respective activity or attrition coefficients (which are normally considered to be constants

independent of the time), and R_x and R_y are their respective reinforcement rates as a function of time into the battle. The functions F and G are usually assumed to not depend explicitly upon the time. For brevity, we sometimes refer to R_x and R_y as the reinforcement schedules. Observe that disengagement or detachment of forces can be treated as negative "reinforcement" rates.

The character of the attrition law is determined by specializing the choice of functions F and G. For example, we have the following:

(1) If $F(x,y) = x$ and $G(x,y) = y$, the resultant attrition law is called the logarithmic law. This name was given to it by Weiss based on work by Peterson [1967]. This attrition law is frequently used, generally in the form of assuming that attrition is determined by a constant (fractional or percentage) loss rate (*i.e.*, so many losses per 1,000 men per day).

(2) If $F(x,y) = y$ and $G(x,y) = x$, the resultant attrition law is called the Lanchester square law. It was apparently proposed independently by both Lanchester and Osipov around 1915, as described in Osipov [1915], Lanchester [1916], Helmbold and Rehm [1995].

(3) If $F(x,y) = xy$ and $G(x,y) = xy$, the resultant attrition law is called the Lanchester linear law. It was proposed by Lanchester [1916].

(4) If $F(x,y) = (xy)^p$ and $G(x,y) = (xy)^p$, the resultant attrition law is called the p-linear law. Lanchester's linear law is the 1-linear law. The 1/2-linear law was proposed by Helmbold [1965]. All p-linear laws with the same attrition coefficients A and D and zero reinforcement rates have exactly the same first integral or state equation, but they attain those states at different times depending on the value of p.

(5) If $F(x,y) = x^{(1-w_1)}y^{w_1}$ and $G(x,y) = y^{(1-w_2)}x^{w_2}$, the resultant attrition law is called the Taylor-Helmbold law. It was proposed by Taylor [1983]. The exponents w_1 and w_2 are called the Weiss parameters. If $w_1 = w_2 = w$, the resultant attrition law has been called the Helmbold law (see Helmbold [1965] and Taylor [1983]). The 1/2-linear law is the special case $w_1 = w_2 = w = 1/2$. The Lanchester square law is the special case $w_1 = w_2 = w = 1$. The logarithmic law is the special case $w_1 = w_2 = w = 0$.

Note that the logarithmic law, the Lanchester square law, the Lanchester linear law, and the 1/2-linear law each involve only two unknown parameters, specifically the attrition coefficients A and D. If we assume that the initial and final strengths, the reinforcement schedules, and the starting and ending times are known exactly for some battle, then these unknown parameters can be uniquely determined by fitting them to the given data. Each of the data items specified must be known if the attrition coefficients are to be uniquely determined.

However, each of the four attrition laws named in the preceding paragraph can be fitted *exactly* to such data! Consequently, an evaluation of the *form* of the attrition law requires some additional information. If we assume all data are known *precisely*, then the strength of the engaged forces at some instant intermediate between the start and end of the battle generally suffices to determine which of the four laws (*i.e.*, logarithmic, square, linear, or 1/2-linear) applies. For historical battles, of course, none of the data are known so accurately (not even the battle's starting and ending times), so estimates of the strengths of the engaged forces at several

points in time intermediate between the start and end of the battle are needed to determine which of the four laws best fits the empirical data.

The p-linear and Helmbold laws each involve three unknown parameters—the two attrition coefficients (A and D) and one exponent (either p or w). If we assume that the initial and final strengths, the reinforcement schedules, and the starting and ending times are known *exactly* for some battle, then this is still not enough information to uniquely determine the values of the three unknown parameters. That is, several possible choices of the three parameters will give equally good fits to the data. The addition of *precise* information about the strengths of the engaged forces at one instant intermediate between the start and end of the battle provides enough information to uniquely determine all three parameters.

However, both the p-linear and Helmbold equations can be fitted *exactly* to such data. Consequently, distinguishing between the Helmbold and p-linear laws requires some additional information. If we assume all data are known *precisely*, then the strength of the engaged forces at a second instant intermediate between the start and end of the battle generally suffices to determine which of the two laws (*i.e.*, p-linear or Helmbold) applies. In practice, none of the data are known so accurately, so estimates of the strengths of the engaged forces at several points in time intermediate between the start and end of the battle are needed to determine which of the two laws best fits the empirical data.

Similarly, the Taylor-Helmbold equations involve four unknown parameters (the two attrition coefficients A and D, and the two exponents w_1 and w_2). To uniquely determine all four of these unknown parameters, the strengths of the engaged forces must be known *exactly* at no less than two instants intermediate between the start and end of the battle (in addition to knowing *exactly* the initial and final strengths, the reinforcement schedules, and the duration of the battle). In practice the strengths must be estimated at several points in time in order to provide a good evaluation of how well the model fits the empirical data.

The Lanchester law has an obvious generalization to heterogeneous forces. This generalization can be written as:

$$\left.\begin{aligned} dx/dt &= -Dy + R_x \\ dy/dy &= -Ax + R_y \end{aligned}\right\}$$

where now x and y are to be interpreted as vectors whose i-th component gives the number of surviving engaged elements of the i-th type at time t after the start of the battle, A and D are matrices of attrition coefficients (normally assumed to be independent of the time), and R_x and R_y are vectors whose i-th components give the reinforcement schedule for the i-th type of combat element. Assume that the reinforcement schedules are known *exactly*. Suppose also that there are m types of combat element for x and n for y. Then A is n-by-m and D is m-by-n, so that in general there are

$$P = 2mn + (m + n)$$

free parameters, $2mn$ for the attrition coefficients and $(m + n)$ for the strengths at the start of the battle. So we need P conditions to determine them uniquely. A complete and *precise* knowledge

of the surviving strengths of all $(m + n)$ types of engaged combat elements at some *precisely-known* instant provides only $(m + n)$ conditions. Hence, to uniquely determine all of the free parameters, we need to have *precise* knowledge of the surviving strengths of all the $(m + n)$ types of combat element at each of at least

$$N = P/(m+n) = 1 + 2mn/(m+n)$$

different points in time. When $m = n$, then $N = n + 1$. For example, when $m = n = 1$ then $N = 2$, which is exactly what we found earlier for the Lanchester square law, for which the strengths at the start and end of the battle sufficed to determine the attrition coefficients. When $m = n = 50$, then $N = 51$, so that *precise* information on the surviving strengths of all 50 + 50 = 100 combat element types for at least 51 different *precisely-known* instants of time is required just to determine the free parameters. Due to the inevitable imprecision of historical data, information on the surviving strengths of all 100 combat element types at a substantially larger number of points in time would actually be needed to decide how well the model fits the data. No one familiar with the kinds of data on historical combat operations will entertain any hope of seeing this much detailed and accurate data in the foreseeable future. However, combat simulations could provide this sort of data, and their results could be compared to those provided by the fitted generalized Lanchester square law equations.

2-4. ARTICLE NUMBER 3

(This article was published in the September 1993 issue of PHALANX.)

Earlier columns gave the necessary and sufficient conditions for Lanchester's square law to hold and discussed the kinds of data needed to test the validity of some simple attrition laws. Here we show how to express the square law and its solution in a form particularly well-suited for further analysis and comparison to empirical data. Since our results as well as notation and terminology will be used in future columns, and space will prevent our restating them, we suggest that readers save this column for future reference.

Start with the familiar square law attrition equations

$$\left.\begin{aligned} x' &= -Dy,\ x(0) = x_0 \\ y' &= -Ax,\ y(0) = y_0 \end{aligned}\right\} \tag{2-1}$$

where A and D are the attrition coefficients, $x(t)$ is the surviving strength of the attacker (mnemonic "attaxxer") and $y(t)$ the surviving strength of the defender (mnemonic "defenyyer") as of time t into the battle, and primes denote differentiation with respect to time. Now, the solution of these equations is well-known. It appears in Prof. James Taylor's [1983] marvelous monograph on Lanchester models of warfare, and at this point in time everything in it can be considered as "well-known."

A physicist looking at these equations might consider replacing its variables with "dimensionless" ones by making the change of variables $a = x(t)/x_0$ and $d = y(t)/y_0$. When this is done equation (2-1) takes the form:

$$\left.\begin{array}{l} a' = -\delta d,\ a(0) = 1 \\ d' = -\alpha a,\ d(0) = 1 \end{array}\right\} \qquad (2\text{-}2)$$

where the new attrition coefficients are given by

$$\left.\begin{array}{l} \alpha = Ax_0 / y_0 = A \div FRY \\ \delta = Dy_0 / x_0 = D \times FRY \end{array}\right\} \qquad (2\text{-}3)$$

where *FRY* stands for the (initial) force ratio favoring the defender, *i.e.*, $FRY = y_0 / x_0$. The usual first integral of equation (2-2) is found by dividing them to eliminate *t*, separating variables and integrating. The result is

$$\mu^2 \equiv \delta / \alpha = \frac{1-a^2}{1-d^2} \qquad (2\text{-}4)$$

where the new quantity μ, introduced as an abbreviation of $\sqrt{\delta / \alpha}$, will turn out to have profound significance.

The solution of (2-2) can be written in the simple and compact form

$$\left.\begin{array}{l} a = \cosh(\lambda t) - \mu \sinh(\lambda t) \\ d = \cosh(\lambda t) - \dfrac{1}{\mu}\sinh(\lambda t) \end{array}\right\} \qquad (2\text{-}5)$$

where $\lambda \equiv \sqrt{\alpha\delta}$, as can readily be confirmed by differentiation. Note that we can convert among all the different attrition coefficients by using the relations:

$$\left.\begin{array}{l} D = \delta \div FRY \\ A = \alpha \times FRY \\ \lambda = \sqrt{\alpha\delta} = \sqrt{AD} \\ \mu = \sqrt{\dfrac{\delta}{\alpha}} = FRY\sqrt{\dfrac{D}{A}} \\ \lambda\mu = \delta \\ \lambda/\mu = \alpha \end{array}\right\} \qquad (2\text{-}6)$$

In future columns we will probe into the physical significance of the coefficients λ and μ, and show that they represent the intensity of combat and the defender's relative advantage over the attacker (respectively). Here we content ourselves with showing that they approximate some other combat parameters with those interpretations. The approximation is motivated by the fact that the fractional losses in real land combat battles seldom exceed 10 percent. This means that in equation (2-5) the quantity λt can be taken as very small compared to unity, so that the hyperbolic functions can be replaced with their approximate values $\cosh(\lambda t) \approx 1$ and $\sinh(\lambda t) \approx \lambda t$. Then equation (2-5) can be written as

$$\left.\begin{aligned} f_x &= 1-a \approx \mu\lambda t \\ f_y &= 1-d \approx \frac{1}{\mu}\lambda t \end{aligned}\right\} \tag{2-7}$$

where f_x and f_y are the fractional losses to the attacker and defender (respectively) and could also be written as

$$\left.\begin{aligned} f_x &= \frac{C_x}{x_0} \\ f_y &= \frac{C_y}{y_0} \end{aligned}\right\} \tag{2-8}$$

where C_x and C_y are the casualties suffered by the attacker and defender (respectively) as of time t into the battle. Multiplying equations (2-7) together we find that

$$\varepsilon \equiv \lambda t \approx \sqrt{f_x f_y} \tag{2-9}$$

the geometric mean of the casualty fractions, which seems to measure reasonably well what historians mean when they speak of a battle as being "bloody," "hard-fought," or "bitter." Accordingly, we will call ε the bitterness parameter. The parameter $\lambda = \varepsilon / t$ is clearly the average rate at which the bitterness is increasing, and so will be called the "intensity" parameter.

Dividing the first of equation (2-7) by the second, we find

$$\mu \approx \sqrt{\frac{f_x}{f_y}} \tag{2-10}$$

so that μ is related to the fractional exchange ratio favoring side Y (the defender), defined by $FERY = f_x/f_y$. When *FERY* is greater than one, the attacker's fractional attrition rate is greater than the defender's. In such a case, the *attacker* is threatened with losing his entire force before the defender does, so the defender has the advantage. When the *FERY* is less than one, the attacker's fractional attrition rate is lower than the defender's. In that case, the *defender* is threatened with losing his entire force before the attacker does, so the attacker has the advantage. Consequently, $\mu > 1$ signals that the defender has the advantage and $\mu < 1$ signals that the attacker has the advantage. Unfortunately, μ is balanced at unity (the point at which neither side has an advantage over the other) and is limited to positive values. A more intuitive measure of side Y's (the defender's) advantage is provided by $ADVY = \ln(\mu)$, the natural logarithm of μ. Now *ADVY* is balanced at zero, is positive when the defender has the advantage, is negative when the attacker has the advantage, and ranges from minus infinity to plus infinity. So we call *ADVY* the defender's relative advantage over the attacker, or (provided it's clear from the context that the defender side is intended) just the advantage parameter for short.

OK. So, at least in a theoretical sense, λ and μ (or *ADVY*) represent the intensity of battle and the defender's advantage, and $\varepsilon = \lambda t$ is a good measure of its bitterness. Do the empirical

data support these interpretations? A future column will take up methods for estimating the mu-parameter (μ), and the advantage parameter $ADVY = \ln(\mu)$ from the empirical data and show that these theoretical interpretations are supported by the data.

2-5. ARTICLE NUMBER 4

(This article was published in the December 1994 issue of PHALANX.)

In a previous column, we started with the familiar square law attrition equations

$$\left.\begin{aligned} x' &= -Dy, x(0) = x_0 \\ y' &= -Ax, y(0) = y_0 \end{aligned}\right\} \tag{2-11}$$

where A and D are the attrition coefficients, x is the surviving strength of the attacker (mnemonic "attaxxer") and y the surviving strength of the defender (mnemonic "defenyyer") as of time t into the battle, and primes denote differentiation with respect to time. We then changed to the dimensionless variables $a = x / x_0$ and $d = y / y_0$ in order to write (2-11) as:

$$\left.\begin{aligned} a' &= -\delta d, a(0) = 1 \\ d' &= -\alpha a, d(0) = 1 \end{aligned}\right\} \tag{2-12}$$

where the new attrition coefficients are given by

$$\left.\begin{aligned} \delta &= Dy_0 / x_0 = D \times FRY \\ \alpha &= Ax_0 / y_0 = A / FRY \end{aligned}\right\} \tag{2-13}$$

where FRY stands for the (initial) force ratio favoring the defender, *i.e.*, $FRY = y_0 / x_0$. We remarked that the usual first integral of equation (2-13), found by dividing them to eliminate t, separating variables and integrating, is

$$\mu^2 \equiv \delta / \alpha = \frac{1-a^2}{1-d^2} \tag{2-14}$$

We also defined the defender's advantage parameter by

$$ADVY = \ln(\mu) \tag{2-15}$$

and remarked that it (or, equivalently, μ) would turn out to have profound significance. ***The remainder of this article explains some of the theoretical basis for that remark.*** In the first place, as pointed out in the previous article, μ is related to the instantaneous fractional exchange ratio favoring the defender via the approximate relation

$$\mu \approx \sqrt{FERY(t)} \equiv \sqrt{f_x / f_y} \tag{2-16}$$

where the instantaneous fractional exchange ratio at time t, $FERY(t)$, is defined by

$$FERY(t) \equiv f_x(t) / f_y(t) \tag{2-17}$$

in which $f_x(t)$ and $f_y(t)$, the instantaneous fractional losses to the attacker and defender (respectively), can be written as

$$\left.\begin{aligned} f_x(t) = C_x(t)/x_0 = 1 - a \\ f_y(t) = C_y(t)/y_0 = 1 - d \end{aligned}\right\} \tag{2-18}$$

where $C_x(t)$ and $C_y(t)$ are the casualties suffered by the attacker and defender (respectively) as of time t into the battle. For future reference, we also define the instantaneous dimensionless force ratio in favor of side Y (the defender) by

$$z(t) = d(t)/a(t) = \frac{y(t)}{x(t)} FRY \tag{2-19}$$

Now, whenever attrition is governed by Lanchester's square law, the defender's advantage parameter $ADVY$ (or μ) can be strongly motivated as an index of defender superiority in battle. This is indicated by the following facts, which are easy consequences of Lanchester's square law and can be used by instructors as simple exercises in Lanchester theory. If $ADVY > 0$ (or $\mu > 1$), then:

a. For all times t after the start of the battle, the defender's instantaneous fractional surviving force exceeds the attacker's, that is, $d(t) > a(t)$. Moreover, their difference in favor of the defender, $d - a$, increases monotonically with time t into the battle in direct proportion to $\sinh(\lambda t)$. Also, the attacker's instantaneous casualty fraction exceeds the defender's, that is, $f_x(t) > f_y(t)$. Moreover, their difference in favor of the defender, $f_x - f_y$, increases monotonically with t in direct proportion to $\sinh(\lambda t)$.

b. The attacker's strength, $x(t)$, reaches zero sooner than the defender's strength, $y(t)$.

c. The instantaneous force ratio favoring the defender, $FRY(t) = y(t)/x(t)$, increases monotonically with time t into the battle. The same is true of the instantaneous dimensionless force ratio, $z(t) = d(t)/a(t)$.

d. The instantaneous casualty exchange ratio favoring the defender, defined as

$$CERY(t) = \frac{C_x(t)}{C_y(t)} = \frac{x_0 - x(t)}{y_0 - y(t)} \tag{2-20}$$

increases monotonically with time t into the battle. The same is true of the instantaneous fractional exchange ratio favoring the defender, defined as

$$FERY(t) \equiv \frac{f_x(t)}{f_y(t)} = \frac{C_x(t)/x_0}{C_y(t)/y_0} = CERY(t) \times FRY = \frac{1-a(t)}{1-d(t)} \tag{2-21}$$

e. The instantaneous differential casualty exchange ratio favoring the defender, defined as

$$DCERY(t) = dx/dy \equiv x'(t)/y'(t) \tag{2-22}$$

increases monotonically with time t into the battle.

f. The instantaneous differential fractional exchange ratio favoring the defender, defined as

$$DFERY(t) = \frac{x'(t)/x(t)}{y'(t)/y(t)} \equiv \frac{a'(t)/a(t)}{d'(t)/d(t)} \tag{2-23}$$

increases monotonically with time t into the battle.

g. Moreover, if *at any* time t during the battle, *either*:

$$\left.\begin{array}{l} d(t) > a(t), \text{or} \\ f_x(t) > f_y(t), \text{or} \\ z(t) > 1, \text{or} \\ FRY(t) > FRY(0) = FRY, \text{or} \\ CERY(t) > FRY, \text{or} \\ FERY(t) > 1, \text{or} \\ DCERY(t) > FRY, \text{or} \\ DFERY(t) > 1 \end{array}\right\} \tag{2-24}$$

then $ADVY > 0$, $\mu > 1$, and *at all* times $t > 0$ during the battle *all* of the following are true:

$$\left.\begin{array}{l} d(t) > a(t), \text{and} \\ f_x(t) > f_y(t), \text{and} \\ z(t) > 1, \text{and} \\ FRY(t) > FRY(0) \equiv FRY, \text{and} \\ CERY(t) > CERY(0) \equiv \mu^2 FRY > FRY, \text{and} \\ FERY(t) > \mu^2 > 1, \text{and} \\ DCERY(t) > DCERY(0) \equiv \mu^2 FRY > FRY, \text{and} \\ DFERY(t) > DFERY(0) \equiv \mu^2 > 1 \end{array}\right\} \tag{2-25}$$

Conversely, when $ADVY < 0$, $\mu < 1$, then all the inequalities in the foregoing propositions are reversed. Hence, at least for the square law attrition equations, the defender's advantage parameter, $ADVY$ (or μ), is an excellent theoretical index or figure of merit for determining

which side has the upper hand. Of course, since it is not clear that real combat follows the square law attrition equations, the degree to which the above propositions are reflected in the behavior of real battles remains an empirical question. We will take up the empirical findings in the next column.

CHAPTER 3

EMPIRICAL JUSTIFICATION

3-1. INTRODUCTION. This chapter presents the six articles (numbers 5 through 10) dealing with the empirical justification for using the advantage parameter as a measure of combat power in practical military applications.

3-2. ARTICLE NUMBER 5

(This article was published in the March 1995 issue of PHALANX.)

In previous columns we identified the defender's advantage parameter, $ADVY = \ln(\mu)$, as a key parameter that, at least according to Lanchester square law theory, should be closely related to victory in battle. In this column we begin to examine the empirical evidence bearing on this matter. However, we open with a quick review of the main theoretical points.

The usual Lanchester square law differential equations are

$$\left.\begin{aligned} x' &= -Dy, x(0) = x_0 \\ y' &= -Ax, y(0) = y_0 \end{aligned}\right\} \qquad (3\text{-}1)$$

where A and D are the attrition coefficients, x and y are the surviving strengths of the attacker (mnemonic "attaxxer") and defender (mnemonic "defendyyer"), respectively, as of time t into the battle, and primes denote differentiation with respect to time. We rewrite equation (3-1) in terms of the "dimensionless" variables $a = x(t)/x_0$ and $d = y(t)/y_0$ as

$$\left.\begin{aligned} a' &= -\delta d, a(0) = 1 \\ d' &= -\alpha a, d(0) = 1 \end{aligned}\right\} \qquad (3\text{-}2)$$

where the new attrition coefficients are given by

$$\left.\begin{aligned} \delta &= Dy_0/x_0 = D \times FRY \\ \alpha &= Ax_0/y_0 = A \div FRY \end{aligned}\right\} \qquad (3\text{-}3)$$

Where $FRY = y_0/x_0$ is the force ratio favoring side Y (the defender). The solution to (3-2) can be written in the form

$$\left.\begin{aligned} a &= \cosh(\lambda t) - \mu \sinh(\lambda t) \\ d &= \cosh(\lambda t) - \frac{1}{\mu}\sinh(\lambda t) \end{aligned}\right\} \qquad (3\text{-}4)$$

Here, we can theoretically identify the constant $\lambda = \sqrt{\alpha\delta} \equiv \sqrt{AD}$ with an index of the intensity of combat, since, by equation (3-1), it is the geometric mean of the instantaneous fractional attrition rates x'/x and y'/y. Similarly, we can theoretically identify $\varepsilon = \lambda T$, where T is the temporal duration of the battle, with an index of the bitterness of the battle. Although it is interesting and informative to study the empirical bases for these theoretical interpretations, our main interest at the moment lies with $\mu = \sqrt{\delta/\alpha}$ and its companion, the advantage parameter favoring the defender, $ADVY = \ln(\mu)$.

Our first task is to find a way of estimating the value of μ from the empirical data usually available on a battle. Now, in the course of solving (3-2) to get (3-4), μ was defined by

$$\mu^2 \equiv \delta / \alpha = \frac{1-a^2}{1-d^2}. \tag{3-5}$$

Now, by definition, $a(T) = \frac{x_0 - C_x(T)}{x_0}$ and $d(T) = \frac{y_0 - C_y(T)}{y_0}$, where $C_x(T)$ and $C_y(T)$ are the casualties taken by the attacker and defender (respectively) during the course of the battle. So, given the initial personnel strengths and the casualties to both sides in a battle, we can determine $a(T)$ and $d(T)$, substitute them into (3-5) to estimate μ for that battle, and use the relationship $ADVY = \ln(\mu)$ to obtain the corresponding advantage parameter favoring the defender, $ADVY$.

Now let's consider the following alternative working hypotheses about victory in combat. In this, we follow Chamberlin's excellent advice [Chamberlin-1965].

H0: Neither the force ratio nor the advantage parameter has an appreciable effect on victory in battle.

H1: Force ratio has an appreciable effect on victory in battle, but the advantage parameter does not.

H3: The advantage parameter has an appreciable effect on victory in battle, but force ratio does not.

H4: Both the force ratio and the advantage parameter have an appreciable effect on victory in battle. (Noted for the sake of completeness, although this hypothesis will not be addressed.)

Quantitative statistical methods have been used to test some of these hypotheses. However, in this exposition, we wish to present simple graphical displays that illustrate the findings. For this purpose, we believe that the most informative displays are smoothed values of the probability of winning a battle versus either the advantage parameter or the logarithm of the force ratio (which we call the logarithmic force ratio, for short). Note that both the advantage parameter and the logarithmic force ratio are "centered," in the sense that zero values theoretically imply that neither side has an edge over its opponent. Also, both the advantage parameter and the logarithmic force ratio are "unbounded," in the sense that each can, at least in principle, vary from minus infinity to plus infinity. Moreover, they are "skew-symmetric," in the sense that, if they have a certain value for a given battle, then interchanging the two sides (that is,

calling the attacker the defender and vice-versa) causes a change in sign of both the advantage parameter and the logarithmic force ratio, but does not alter their absolute magnitudes. Consequently, we might expect our empirical graphs to exhibit the same properties of centeredness, unboundedness, and skew-symmetry.

The graph presented here use data on 83 battles, ranging in date from 1741 to 1945, taken from [Helmbold-1961], also given in [Helmbold-1991]. For each of these battles, this data base records whether the defender won ($WINY = 1$) or lost ($WINY = 0$). However, for our purposes we need an estimate of the *probability* that the defender wins the battle, $\Pr(WINY = 1)$. Some form of smoothing is commonly used to convert qualitative (win/lose, $WINY = 0$ or 1) to quantitative (probability) information, and we cheerfully adopt this method for our computations. In particular, the results presented here smooth the data using a Gaussian kernel with standard deviation equation to 1/10 (where the independent variables are either the advantage parameter favoring the defender or the logarithmic force ratio favoring the defender). Appendix D contains some introductory material on kernel smoothing.

For example, if a battle had an advantage parameter favoring the defender of 0.2, we took that battle weighted with all others according to a Gaussian kernel centered at 0.2. We then used this weighted average proportion of times the defender won as our estimate of the probability that the defender would win a battle whose advantage parameter favoring the defender was equal to 0.2. Using other smoothing methods does not change the main conclusions stated below.

Figure 3-1 displays the smoothed probability that the defender wins versus the advantage parameter favoring side Y (the defender), *ADVY.* As can be seen, the probability that the defender wins does indeed depend significantly on the value of the advantage parameter favoring the defender, *ADVY.* The curve is centered in the sense that $\Pr(WINY) = 0.50$ when *ADVY* is approximately zero. Moreover, the fitted curve tends to be skew-symmetric about the 50 percent probability of winning level. When *ADVY* is used to classify the battles in this data base according to which side won, it correctly classifies about 70 percent of them. Of course, this classification procedure makes no *direct* use of any information on what year the battle was fought, what type or size of forces were engaged on either side, what the losses were on either side, how long the battle lasted, how good the leadership or intelligence was on either side, what tactical plan was adopted by the contending sides, what the logistical and larger strategic situation was on either side, or what the weather or terrain was like. Whatever influence these factors may or may not have exerted on the outcome are condensed into the advantage parameter favoring the defender, from that into the binary classification $ADVY > 0$ or $ADVY < 0$, and from that into a simple binary output of "defender more (or less) likely to win than the defender." This binary output is the only value used *directly* in this classification procedure.

Figure 3-1 also displays the smoothed probability that the defender wins plotted against the force ratio favoring the defender, defined as y_0 / x_0 and shown on a logarithmic scale. These data were also smoothed using a Gaussian kernel with standard deviation equal to 1/10. As can be seen, the probability that the defender wins is not very sensitive to the value of the force ratio favoring the defender. If the sign of the logarithmic force ratio were used to classify the battles in this data base by which side won, it would correctly classify only about 57 percent of them. Since about 52 percent of the battles in this data base were won by the attacker, a 52 percent successful classification would result simply from classifying them *all* as attacker wins. A 57

percent correct classification is hardly any improvement over simply classifying all the battles as attacker wins.

These results provide substantial empirical evidence that the advantage parameter favoring the defender has an appreciable influence on which side wins a battle. Both theoretically and empirically, it appears to capture the essence of those factors most conducive to victory in battle. On the other hand, the force ratio favoring the defender has little or no influence on which side wins. In a future column we will show that these results are not peculiar to the SP-128 data base used here, but are confirmed by other data bases.

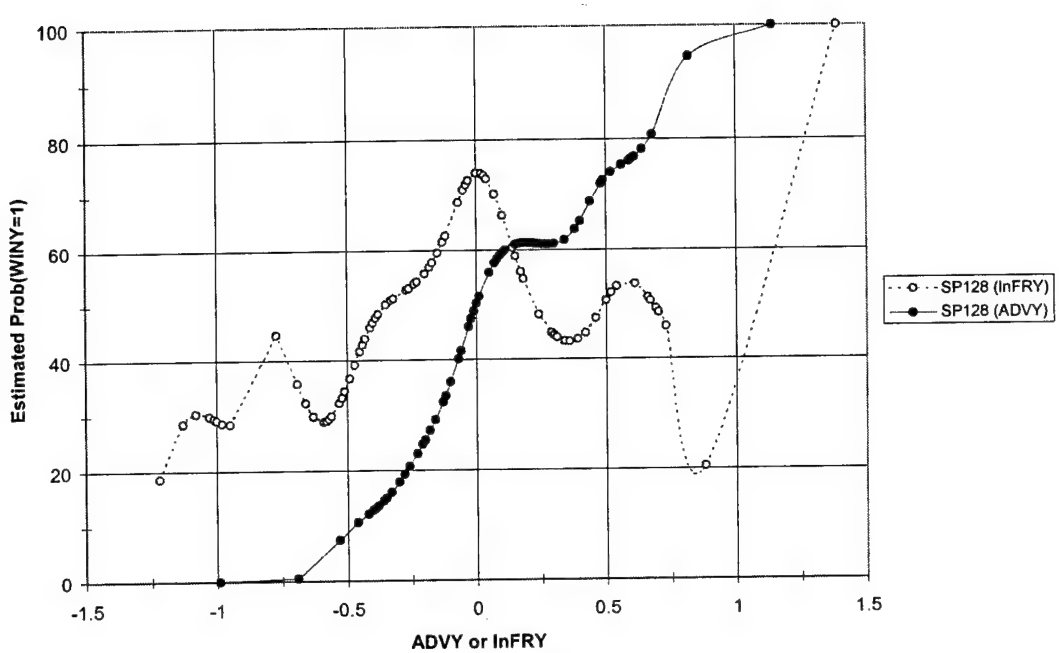

Figure 3-1. Smoothed Estimates of Probability of Winning versus *ADVY* or ln(*FRY*) for the SP-128 Data

3-3. ARTICLE NUMBER 6

(This article was published in the March 1996 issue of PHALANX.)

In previous columns we identified the defender's advantage parameter, $ADVY = \ln(\mu)$ as a key parameter that, at least according to Lanchester square law theory, should be closely related to victory in battle. In the March 1995 issue of the PHALANX (see paragraph 3-2), we presented some empirical evidence in support of this proposition. The reader is referred to that column for the technical definition of the defender's advantage parameter and other essential mathematical background. In this column we continue our examination of the empirical evidence bearing on this matter.

As explained in our March 1995 column (see paragraph 3-2), we wish to present simple graphical displays that illustrate the findings. For this purpose, we believe that the most informative displays are smoothed values of the probability of winning a battle versus either the advantage parameter or the logarithm of the force ratio (which we call the logarithmic force ratio, for short).

Figure 3-2 was prepared using data in the US Army Concepts Analysis Agency data base of battles (CDB90DAT), which contains data on a total of 660 battles from 1600 AD to recent times. This data base was originally created for the Concepts Analysis Agency by Trevor Dupuy and his staff of HERO military historians. The Concepts Analysis Agency then led a series of efforts by HERO and others to further enhance its accuracy and completeness. Despite any remaining faults and limitations, it is the largest, most reliable, and detailed data base of battles currently available. It can be obtained from DTIC or NTIS as *Database of Battles–Version 1990 (Computer Diskette),* US Army Concepts Analysis Agency Data Base, 30 April 1991, AD-M000 121. (See also [Helmbold-1993].)

For our purposes, we will use only those battles that had no reinforcements to either side. This criterion was satisfied by 290 battles of the CDB90DAT data base. We smoothed the data on defender wins using a Gaussian kernel with standard deviation equal to 0.1. Appendix D contains some introductory material on kernel smoothing.

Figure 3-2 displays the smoothed probability that the defender wins versus the advantage parameter favoring the defender, *ADVY.* It also shows the smoothed probability that the defender wins versus the logarithmic force ratio favoring the defender (labeled ln*FRY* on the figure).

When the *ADVY* value is used to classify battles in this data base according to which side won, it correctly classifies about 83 percent of them. Of course, this classification procedure makes no *direct* use of any information on what year the battle was fought, what type or size of forces were engaged on either side, what the losses were on either side, how long the battle lasted, how good the leadership or intelligence was on either side, what tactical plan was adopted by the contending sides, what the logistical and larger strategic situation was on either side, or what the weather or terrain was like. Whatever influence these factors may or may not have exerted on the outcome are condensed into the advantage parameter favoring the defender, and from that into a simple binary output of "defender more (or less) likely to win than the attacker." This binary output is the only value used *directly* in the classification procedure.

When the logarithmic force ratio is used to classify battles in this data base according to which side won, it correctly classifies about 54±3 percent of them. Here the error bounds are the conventional standard errors obtained by assuming that the number correctly classified follows a binomial distribution. Since in this data base only about 39 percent of the battles were won by the defender (and about 61 percent by the attacker), simply classifying *all* battles in this data base as attacker wins would result in correctly classifying about 61 percent of them. So using the logarithmic force ratio to classify battles according to which side won is somewhat *less* successful than simply classifying *all* of them as attacker wins.

Note, too, that the advantage parameter favoring the defender (*ADVY*) is often unambiguous about which side wins. For example, *ADVY* values over 0.25 in absolute value correspond to

probabilities over 80 percent. By contrast, the logarithmic force ratio is almost always ambiguous. That is, its estimate of the probability tends to be around 40 percent independently of the value of the logarithmic force ratio, and does not attain higher probability values even for logarithmic force ratios approaching 1.25. Since a logarithmic force ratio of 1.25 corresponds to a force ratio of about 3.5 to 1, the data indicate that even for these high force ratios the logarithmic force ratio is an ambiguous indicator of which side will win.

These results provide substantial empirical evidence that the advantage parameter favoring the defender has an appreciable influence on which side wins a battle and that the logarithmic force ratio does not. Both theoretically and empirically, the advantage parameter appears to capture the essence of those factors most conducive to victory in battle. On the other hand, the force ratio favoring the defender has little or no influence on which side wins. In a future column we will show that these results are not peculiar to the data bases so far mentioned in this column, but are confirmed by other data bases.

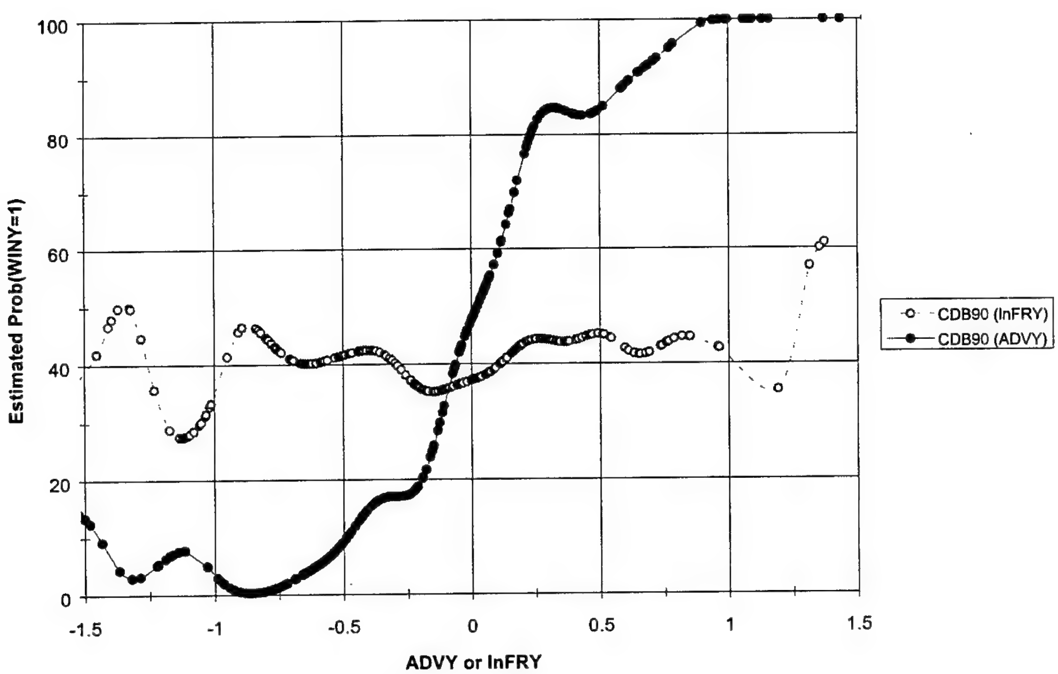

Figure 3-2. Smoothed Estimates of Probability of Winning versus *ADVY* or ln(*FRY*) for the CDB90 Data

3-4. ARTICLE NUMBER 7

(This article was published in the June 1996 issue of PHALANX.)

In previous columns we identified the defender's advantage parameter, $ADVY = \ln(\mu)$, as a key parameter that, at least according to Lanchester square law theory, should be closely related to victory in battle. In the March 1995 (see paragraph 3-2) and September 1995 (see paragraph 3-3) issues of PHALANX, we presented some empirical evidence in support of this proposition. The reader is referred to the March 1995 column (see paragraph 3-2) for the technical definition

of the defender's advantage parameter and other essential mathematical background. In this column we continue our examination of the empirical evidence bearing on this matter.

As explained in our previous columns, we wish to present simple graphical displays that illustrate the findings. For this purpose, we believe that the most informative displays are smoothed values of the probability of winning a battle versus either the advantage parameter or the logarithm of the force ratio (which we call the logarithmic force ratio, for short).

The accompanying graph was prepared using data in the US Army Concepts Analysis Agency's PARCOMBO data base of battles. It contains data on a total of 367 battles that occurred from 280 BC to 1965 AD. This data base was created for the Concepts Analysis Agency by Dr. Robert L. Helmbold as a composite from other data bases. The quality and reliability of the data used in compiling this data base is uneven. Some of the data used are of fairly high quality, while others are of noticeably lower quality. On the average, the PARCOMBO data are considered to be less reliable than those used in previous columns. However, the period of time covered—and hence the variety of battle experience represented—is much greater. All of the data bases used in composing the PARCOMBO data base are obtainable from DTIC or NTIS (see [Helmbold-1993]).

In this column, we will use only those 357 battles from the PARCOMBO data base that have enough data to be usable for our present purposes. That means they have complete data on the strengths and losses of both sides, together with an identification of the winning side. We smoothed the data on defender wins using a Gaussian kernel with standard deviation equal to 0.1. Appendix D contains some introductory material on kernel smoothing.

Figure 3-3 displays the smoothed probability that the defender wins versus the advantage parameter favoring the defender, *ADVY*. It also shows the smoothed probability that the defender wins versus the logarithmic force ratio favoring the defender (labeled ln*FRY* on the figure).

When the *ADVY* value is used to classify battles in this data base according to which side won, it correctly classifies about 77 ± 2 percent of them. Of course, this classification procedure makes no *direct* use of any information on what year the battle was fought, what type or size of forces were engaged on either side, what the losses were on either side, how long the battle lasted, how good the leadership or intelligence was on either side, what tactical plan was adopted by the contending sides, what the logistical and larger strategic situation was on either side, or what the weather or terrain was like. Whatever influence these factors may or may not have exerted on the outcome are condensed into the advantage parameter favoring the defender, and from that into a simple binary output of "defender more (or less) likely to win than the attacker." This binary output is the only value used *directly* in the classification procedure.

When the logarithmic force ratio is used to classify battles in this data base according to which side won, it correctly classifies about 52 ± 3 percent of them. Here the error bounds are the conventional standard errors obtained by assuming that the number correctly classified follows a binomial distribution. Since in this data base about 65 percent of the battles were won by the attacker, simply classifying *all* battles in this data base as attacker wins would result in correctly classifying about 65 percent of them. So using the logarithmic force ratio to classify battles

according to which side won is decidedly *less* successful than simply classifying *all* of them as attacker wins.

Note, too, that the advantage parameter favoring the defender (*ADVY*) is often unambiguous about which side wins. For example, *ADVY* values over 0.25 in absolute value correspond to probabilities in excess of about 60 to 80 percent. By contrast, the logarithmic force ratio is usually ambiguous. That is, its estimate of the probability tends to be around 30 to 40 percent independently of the value of the logarithmic force ratio.

These results, especially when considered in light of the other empirical findings presented in these columns, provide substantial evidence that the advantage parameter favoring the defender has an appreciable influence on which side wins a battle. Both theoretically and empirically, it appears to capture the essence of those factors most conducive to victory in battle. On the other hand, the force ratio favoring the defender has little or no influence on which side wins. In a future column we will show that these results are not peculiar to the data bases so far mentioned in this column, but are confirmed by other data bases.

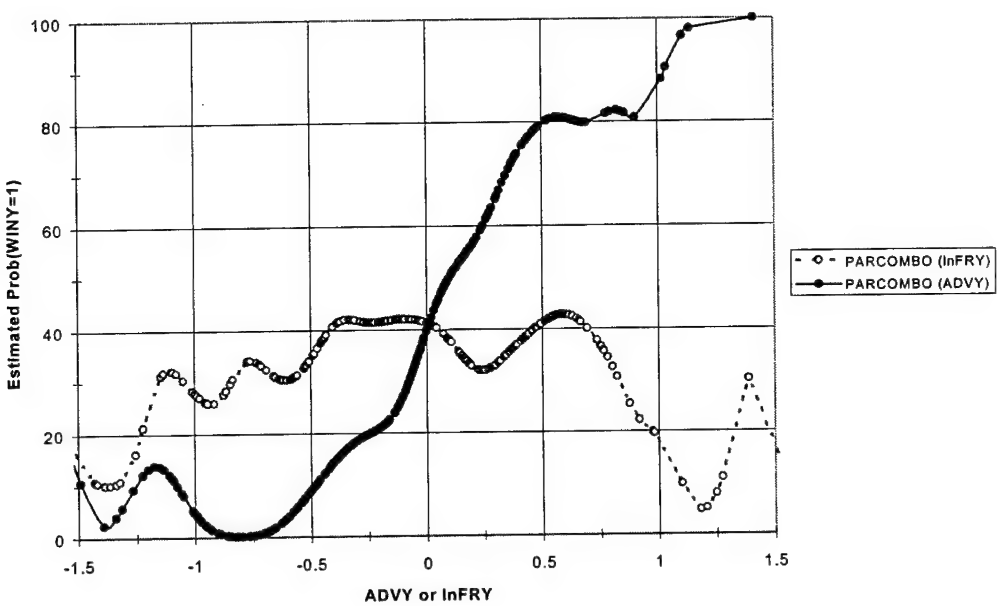

Figure 3-3. Smoothed Estimates of Probability of Winning versus *ADVY* or ln(*FRY*) for the PARCOMBO Data

3-5. ARTICLE NUMBER 8

(This article was published in the December 1996 issue of PHALANX.)

Well, lets start off by reviewing where we are in this series on the defender's advantage parameter, $ADVY = \ln(\mu)$, as a measure of effectiveness in combat. Previous PHALANX Combat Analysis columns have taken up the following topics. In March 1993 (see paragraph 2-3) we gave necessary and sufficient conditions for the validity of Lanchester's square law. In September 1993 (see paragraph 2-4) we presented our version of the solution of Lanchester's

square law, and showed how it naturally led to consideration of the lambda (λ) and mu (μ) parameters. We also noted that these are good indexes of the intensity and defender's advantage, respectively. In December 1994 (see paragraph 2-5), we showed in detail that the advantage parameter governs a remarkably wide variety of qualitative properties related to possession of the advantage in combat operations. All of these results were based on theoretical considerations.

Our first column testing this theory against historical combat data appeared in March 1995 (see paragraph 3-2). It used the SP-128 data base to show that the advantage parameter is empirically a much better predictor of victory in battle than the logarithmic force ratio (the logarithm of the force ratio). In March 1996 (see paragraph 3-3), we showed that this is also true for the much larger and more accurate CDB90DAT data base. In June 1996 (see paragraph 3-4), we showed that it is also true for the PARCOMBO data base, which is more varied than CDB90DAT but less accurate.

This column continues this series of empirical tests using different data bases. Here we use data from Bodart's massive dictionary of battles [Bodart-1908]. Bodart gives both side's strengths and losses for over 1,000 land and sea battles that took place between 1618 and 1905. Bodart also indicates which side won the battle. Unfortunately, Bodart does not say which side was the attacker and which the defender. So we lack the information needed to prepare graphical displays of the kind we used in earlier empirical tests and are forced to use tables instead of graphs. All of Bodart's numerically complete data on battles have been reduced to electronic data base format for convenience in performing analyses [see Helmbold-1993].

We first consider Bodart's land combat data base in electronic form (the BWSH data base). It has 1,087 entries for battles that occurred between 1619 and 1905, inclusive. We use the *ADVY* parameter to predict which side won. This is done by computing from Bodart's strength and loss data the value of the *ADVY* parameter favoring one of the sides (side Y). If the *ADVY* parameter favoring that side is greater than zero, then we "guess" that Bodart will identify side Y as the winner; otherwise we "guess" that Bodart will identify that side Y as the loser. We do the analogous thing, using force ratios (*FRY*s) greater than one as the basis for guessing the winner. We then ask how often these guesses are correct. Some of Bodart's battles are sieges, which we treat separately because the theory that predicts victory based on *ADVY* may apply better to nonsiege battles than to sieges. Table 3-1 gives the results. One of the 1,087 battles lacked the information needed to compute the *ADVY* parameter, so only 1,086 battles are used here. The *ADVY* parameter correctly predicted the winner in 946 (87 percent) of these battles. The *FRY* parameter got 718 (66 percent) of them right. When we deleted the sieges, we were left with 973 battles. The *ADVY* parameter correctly predicted the winner in 845 (87 percent) of these battles; the *FRY* parameter got 626 (64 percent) of them right.

Table 3-1. Percentage of Bodart's Land Battle Winners Correctly Predicted by the *ADVY* or *FRY* Parameters

Data base	Number correct based on *ADVY*	Out of total	Percent correct based on *ADVY*	Number correct based on *FRY*	Percent correct based on *FRY*
BWSH-All	946	1,086	87%	718	66%
BWSH-NoSiege	845	973	87%	626	64%

We now turn to Bodart's sea battles. The BODASHIP data base gives data on 120 naval battles that occurred between 1638 and 1905, inclusive. As for the land battles, Bodart identifies the winning side, but does not say which side was the attacker and which the defender. Now, for naval battles, one might argue that the *ADVY* parameter computation ought to be based on the numbers of ships present and sunk on each side. However, we will not take that approach. Instead, even for naval battles, we continue to base the *ADVY* parameter on the personnel strengths and casualties on each side. This is consistent with what was done for land battles. It also avoids the very difficult task of seeking an index that adequately rates the comparative "fighting strength" of different types and sizes of ships. Those who would rather base sea battle *ADVY* parameters on the numbers of ships can consider that our approach results in a particularly stringent test of the theory. Table 3-2 gives the results. Only 96 of Bodart's sea battles have enough information to compute the *ADVY* parameter, but 105 of them have enough data to compute the (personnel) force ratio. The *ADVY* parameter correctly predicts the winner in 86 out of 96 battles (90 percent). The force ratio got 56 out of 105 battles (53 percent) right.

Table 3-2. Percentage of Bodart's Sea Battle Winners Correctly Predicted by the *ADVY* Parameter

Data base	Number correct based on *ADVY*	Out of total	Percent correct based on *ADVY*	Number correct based on *FRY*	Percent correct based on *FRY*
BODASHIP	86	96	90%	56	53%

So we have now shown that the *ADVY* parameter is quite successful at predicting victory in battle for the SP-128, CDB90DAT, PARCOMBO, BWSH, and BODASHIP data bases. In a future column we will discuss its application to a data base of wars, as distinguished from battles.

3-6. ARTICLE NUMBER 9

(This article was published in the March 1997 issue of PHALANX.)

As usual, be begin by reviewing where we are in this series on the defender's advantage parameter, $ADVY = \ln(\mu)$, as a measure of effectiveness in combat. Previous PHALANX Combat Analysis columns took up the following topics. In March 1993 (see paragraph 2-3) we gave necessary and sufficient conditions for the validity of Lanchester's square law. In September 1993 (see paragraph 2-4) we presented our version of the solution of Lanchester's square law, and showed how it naturally led to consideration of the lambda (λ) and mu (μ) parameters. We also noted that these are good indexes of the intensity and defender's advantage, respectively. In December 1994 (see paragraph 2-5), we showed in detail that the advantage parameter governs a remarkably wide variety of qualitative properties related to possession of the advantage in combat operations. However, all of the preceding results are based on purely theoretical considerations.

Our first column testing this theory against historical combat data appeared in March 1995 (see paragraph 3-2). It used the SP-128 data base to show that the advantage parameter is empirically a much better predictor of victory in battle than the logarithmic force ratio (the logarithm of the force ratio). In March 1996 (see paragraph 3-3), we showed that this is also true for the much larger and more accurate CDB90DAT data base. A test showing that it is also true for the PARCOMBO data base, which is more varied than CDB90DAT but less accurate, appeared in the June 1996 issue of PHALANX (see paragraph 3-4). A test based on Bodart's massive dictionary of battles appeared in the December 1996 issue (see paragraph 3-5). These columns showed that the *ADVY* parameter is quite successful at predicting victory in land combat battles for the SP-128, CDB90DAT, PARCOMBO, and BWSH land battle data bases. The *ADVY* parameter was also shown to be quite successful at predicting the victor for the naval battles given in Bodart's dictionary of battles (the BODASHIP data base).

In this column we consider Small's data [Small-1982] on 67 international wars that began between 1823 and 1980. The average and standard deviation of the war dates is around 1907 and 44, respectively. Three of these 67 wars were still on-going at the time Small wrote, so only 64 have final data. We limit our attention to the 64 with final data. Small also rated his data as "high" and "low" to reflect his relative degree of confidence them (although his pages 73 and 74 make clear that essentially all these data are approximate). There are 29 "high" confidence wars, with dates ranging from 1823 to 1973. The average and standard deviation of the dates of these 29 wars is 1894 and 40, respectively.

Our analysis of these data generally follows [Helmbold-1987]. In interpreting the results, several differences between the war and the battle data need to be kept in mind. The most important of these are described below. First, we elected to use mainly Small's 29 "high" confidence wars. Secondly, in our previous work on battles we had data on the attacker's and defender's strength and battle casualties. Small does not give these values for wars. We had to assume that the side described by Small as the war's "initiator" corresponded to the attacker in a battle, even though the sides often reversed roles during a war. It was also necessary to use Small's determination of pre-war national population as the nearest correspondent to the initial combat personnel strength of the opposing forces in a battle, and to use a value characterized by Small as "combat-related deaths of military personnel" as a surrogate for battle casualties.

Secondly, in our previous work on *battles* we calculated the *ADVY* parameter from data on the opposing force strengths and their losses. For *wars*, we had to estimate the *ADVY* parameter from Small's value for the ratio of the opponent's to the initiator's combat-related deaths as a fraction of their respective pre-war populations, since that is the only value Small provides. Small's ratio can be expressed algebraically as $S = I / O$, where I is the ratio of the initiator's combat-related battle deaths to the initiator's pre-war population and O is the corresponding value for the opponent. We interpreted Small's ratio as roughly corresponding to the fractional exchange ratio favoring the opponent (*i.e.*, $S \approx FERY$), and estimated the corresponding *ADVY* parameter favoring the initiator as half the natural logarithm of that fractional exchange ratio (*i.e.*, $ADVY \approx \frac{1}{2}\ln(FERY) \approx \frac{1}{2}\ln(S)$). A justification for estimating the *ADVY* parameter as half the logarithm of the fractional exchange ratio is provided in the Analyst's Corner column of the September 1993 issue of PHALANX (see equation (2-10) in paragraph 2-4).

Thirdly, for battles we used wherever possible only those which had few or no reinforcements, so that their initial strength could be confidently used as the basis for percentage losses. This is obviously not possible with the war data. Fourthly, we made no allowance for possible inaccuracies in the data, and instead took Small's reported data at face value.

Further differences between battles and wars are due to the facts that (i) the use of the entire national population as a surrogate for personnel strength is questionable because infants, the elderly, the institutionalized, and (often) women are not subject to combat duty (so it would be better to use instead the population potentially subject to combat duty—were such data available), (ii) even in the best of circumstances the use of just the personnel strengths and losses may inadequately represent the complex ebb and flow of historical events during an entire war, and (iii) political, economic, sociological, and strategic considerations generally have much greater influence on wars than on battles.

In view of these rough approximations, it would plainly be at least a minor miracle to find that the relation of the *ADVY* parameter to victory in these wars is about the same as previously found for battles. However, Figure 3-4 shows that for the "high" confidence war data that is in fact what occurs. Because there are only 29 high confidence wars, these data were smoothed using a Gaussian kernel with a standard deviation of 0.15 rather than 0.10. The estimated *ADVY* parameter is able to predict victory in about 79 percent of the cases when only the 29 "high" confidence wars are considered, and in about 72 percent of the cases when all 64 wars are considered. (These predictions are based solely on the algebraic sign of the *ADVY* parameter—somewhat better predictions could be obtained by taking into consideration its magnitude as well as its sign).

To sum up, we have now shown that from an empirical as well as from a theoretical standpoint the *ADVY* parameter is a reasonable quantitative index for the probability or "decisiveness" of victory in wars as well as in land and naval battles. Accordingly, it can with a reasonable degree of confidence be used as an empirically-based numerical measure of the overall effectiveness or "combat power" of a force *vis-à-vis* its opponent when evaluating alternative force structures, operational plans, procurement programs, and in other military analysis situations.

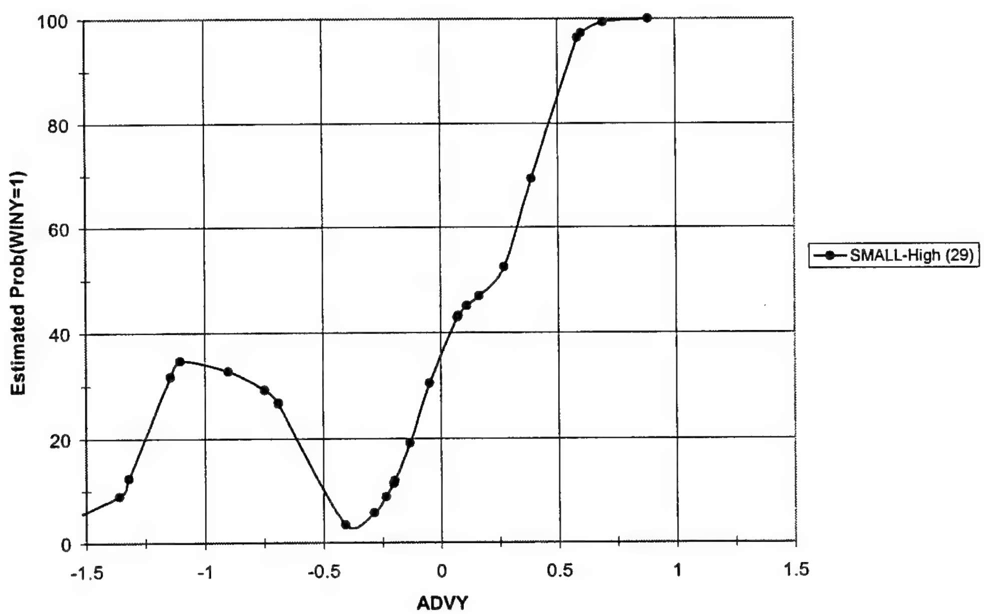

Figure 3-4. Smoothed Estimate of Probability of Winning versus *ADVY* for Small's High Confidence War Data

3-7. ARTICLE NUMBER 10

(This article was submitted to PHALANX in January 1997.)

We now come to the close of this series of articles on the defender's advantage parameter, $ADVY = \ln(\mu)$, as a measure of effectiveness in combat. This series began in December 1992

(see paragraph 2-2) and in March 1993 (see paragraph 2-3), when we discussed the necessary and sufficient conditions for the validity of Lanchester's square law. In September 1993 (see paragraph 2-4) we presented our version of the solution of Lanchester's square law, and showed how it naturally led to consideration of the lambda (λ) and mu (μ) parameters. We also noted that these are good indexes of the intensity and defender's advantage, respectively. In December 1994 (see paragraph 2-5), we showed in detail that the advantage parameter governs a remarkably wide variety of qualitative properties related to possession of the advantage in combat operations. However, all of the preceding results are based on purely theoretical considerations.

Our first column testing this theory against historical combat data appeared in March 1995 (see paragraph 3-2). It used the SP-128 data base to show that the advantage parameter is empirically a much better predictor of victory in battle than the logarithmic force ratio (the logarithm of the force ratio). In March 1996 (see paragraph 3-3), we showed that this is also true for the much larger and more accurate CDB90DAT data base. A test showing that it is also true for the PARCOMBO data base, which is more varied than CDB90DAT but less accurate, appeared in the June 1996 issue of PHALANX (see paragraph 3-4). A test based on Bodart's massive dictionary of battles appeared in the December 1996 issue (see paragraph 3-5). These columns showed that the *ADVY* parameter is quite successful at predicting victory in land combat battles for the SP-128, CDB90DAT, PARCOMBO, and Bodart data bases. The *ADVY* parameter was also shown to be quite successful at predicting the victor for the naval battles given in Bodart's dictionary of battles. An article showing that these results are valid for wars as well as battles was submitted for publication in September 1996 (see paragraph 3-6).

In this column we provide a graph (see Figure 3-5 below) overlaying the results for all the data bases used to compute a smoothed probability of winning versus the defender's advantage parameter. For each data base, it shows the smoothed probability that the defender wins, Pr(*WINY*), as a function of the defender's advantage parameter, *ADVY*. As can be seen, each data base shows that the probability of winning generally tends to increase as the defender's advantage parameter increases.

This has enormous practical value for military operations. It teaches us that victory in battles and wars depends on obtaining a large favorable advantage parameter. Accordingly, combat operations should be conducted with the aim of maximizing the favorable advantage parameter. On both theoretical and empirical grounds, this can be accomplished by maximizing the favorable fractional exchange ratio. The fractional exchange ratio favorable to the defender is defined as

$$FERY = \frac{\text{Percentage attrition to the attacker}}{\text{Percentage attrition to the defender}}$$

The fractional exchange ratio favoring the attacker, $FERX$ is just the reciprocal of this, so that $FERX \equiv 1/FERY$. For example, if at some point in time the attacker and defender have taken losses amounting to 3 percent and 2 percent, respectively, then $FERY = 3/2 = 1.5$ and $FERX = 2/3 = 0.667$. The corresponding value of the defender's advantage parameter will be close to $ADVY \cong \frac{1}{2}\ln(FERY) = 0.20$. If an alternative course of action would change the losses to 4 percent and 2 percent for the attacker and defender, respectively, then the fractional exchange

ratio favoring the defender would increase to $FERY = 2$ and the advantage parameter favoring the defender would increase to about 0.35. Accordingly, the alternative course of action would normally be preferred by the defender.

When, as in Figure 3-5, all the curves relating probability of winning to the advantage parameter are overlaid, we note that there seems to be a fairly consistent tendency for the curves to have a “bump” or peculiar rise near $ADVY = -0.75$. This is a new phenomenon. It was discovered only when these curves were overlaid. At present, no one has an explanation for this. Gaining more insight as to what is causing it is a basic research problem worthy of further investigation. Various possibilities could be considered. For example, one speculation is that there are few battles in this region, so the curves are easily confused by a very small number of unusual or erroneous data points. Another speculation is that the bump may merely be due to the data bases containing an unusually large fraction of dramatic and memorable instances of the defender “snatching victory from the jaws of defeat,” because these tend to be more carefully recorded and intensively studied than the normal cases. Yet another speculation is that the bump may represent cases where the defender had no satisfactory means of withdrawing or breaking contact, and with his “back to the wall” was faced with the choice of either surrendering the whole force or continuing to fight a basically losing battle—either in the hope of eventual reinforcement and relief or because the mission called for a “last ditch” defense. Perhaps the bump represents cases where the attacker could not learn the extent of the defender’s losses, and so prematurely abandoned the attack without realizing just how desperate the defender’s situation had become. Or some other, as yet unimagined but potentially important, mechanism may be at work.

As mentioned earlier, this closes the series of articles on the defender’s advantage parameter and its use as a measure of effectiveness in combat. We have presented all of the data on this subject that we have available.

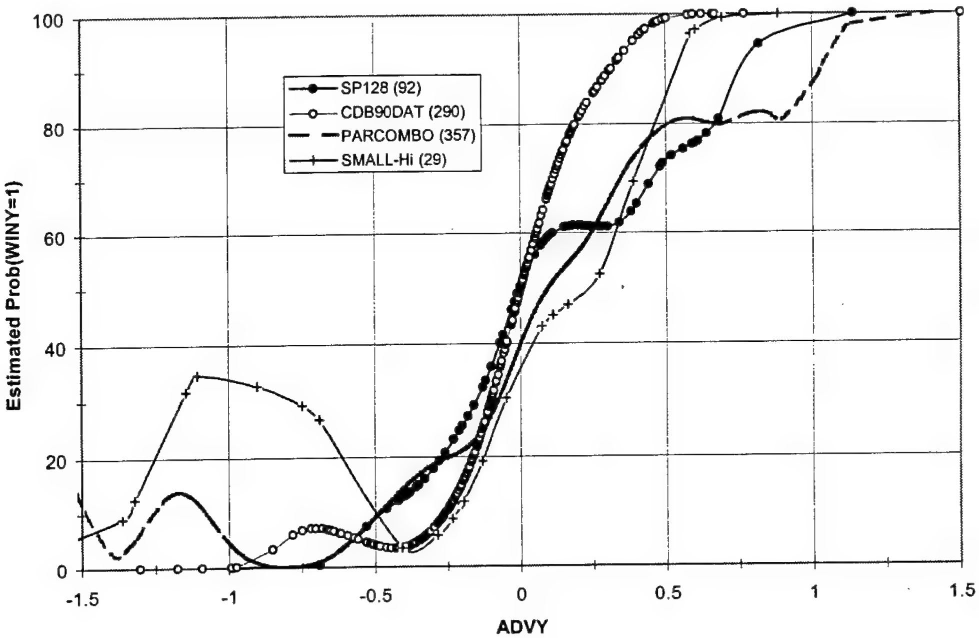

Figure 3-5. Composite Overlay of Probability of Winning versus the Advantage Parameter

APPENDIX A

CONTRIBUTORS

A-1. TEAM

a. Research Director

Dr. Robert L. Helmbold, Tactical Analysis Division

A-2. PRODUCT REVIEW BOARD

Mr. Ronald J. Iekel, Chairman

APPENDIX B

REQUEST FOR ANALYTICAL SUPPORT

PART 1

REQUEST FOR ANALYTICAL SUPPORT

1. Performing Directorate/ Division: TA	**2. Account Number:** 97073
3. Type Effort (Enter one): P (S - Study; Q - QRA; P - Project; R - RAA; M - MMS) Mode (Contract=C):	**4. Tasking (Enter one):** V (F - Formal Directive; I - Informal; V - Verbal)

5. Title: Prepare Memorandum Report documenting PHALANX articles

6. Acronym: ADVReport	**7. Date Request Received:** 01/15/97	**8. Date Due:** 06/30/97

9. Requester/Sponsor (i.e., DCSOPS): Director	**10. Sponsor Division (i.e., SSW, N/A)** ZA

11. Impact on Other Studies, QRA, Projects, RAA: Delay PAR-Phase 5 for 2-3 months

12. Product Required: Memorandum Report

13. Estimated Resources Required:	a. Estimated PSM: 3.0	b. Estimated Funds: $0
c. Models Req'd: None		d. Other: None

14. Objective(s)/Abstract:
Prepare CAA Memorandum Report documenting the series of PHALANX articles on the advantage parameter.

15. Study Director/POC:	Last Name: Helmbold	First: Robert	Date: 01/17/97
	Signature: Robert Helmbold		Phone#: 295-5278

GO TO BLOCK 20 *If this is A STUDY*. See Tab C of the Study Directors' Guide for preparation of a Formal Study Directive.

PART 2

16. Background/Statement of Problem*: Director verbally requested this.

17. Scope of Work*: Prepare CAA Memorandum Report documenting the series of PHALANX articles on the advantage parameter.

18. Issues for Analysis*: None

19. Milestones/Plan of Action*: None

20. Division Chief Concurrence:	[signature]	Date: 2/16/97
21. Sponsor (COL/DA Div Chief) Concurrence:		Date:

22. Sponsor Comments*:

CAA Form 233 (1 May 95)   * Continue on separate sheet   Previous editions Obsolete

APPENDIX C

BIBLIOGRAPHY

Bodart, Gaston, *Militar-Historisches Kriegs-Lexicon (1618-1905)*, C. W. Stern, Vienna and Leipzig, 1908, 956 pp, UNCLASSIFIED, available from libraries.

Busse, James, J., "An Attempt to Verify Lanchester's Equations," Marine Corps Operations Analysis Group (MCOAG), of the Center for Naval Analyses, Unnumbered, circa 1968, 17 pp, UNCLASSIFIED. Essentially the same paper was published as paper 9-1 in "Developments in Operations Research, Vol 2," Proceedings of the Third Annual Israel Conference on Operations Research, held at Tel-Aviv, Israel, July 1969, Benjamin Avi-Itzhak (ed.), Gordon and Breach Science Publishers, New York, 1971, ISBN: 0-677-39840-X, which is available from libraries.

Chamberlin, T. C., The Method of Multiple Working Hypotheses, Science, vol 148 (1965), pp 754-759, UNCLASSIFIED, available from libraries.

Dunn, L. P., "A Lanchester Fit to Selected Land Battles," Lulejian and Associates, February 1971, 46 pp, UNCLASSIFIED. May be available from DTIC. A copy of the original is in Helmbold's personal files. Prepared for Assistant Chief of Staff, Studies and Analysis, Headquarters, US Air Force, under Contract No. F44620-70-C-0056.

Engel, Joseph H., "A Verification of Lanchester's Law," Journal of the Operations Research Society of America, Vol 2, No 2, May, 1954, pp 163-171. UNCLASSIFIED, available from libraries.

Fain, William W.; Fain, Janice B.; Feldman, Leon; and Simon, Susan, "Validation of Combat Models Against Historical Data," Consolidated Analysis Centers Inc. (CACI) and Center for Naval Analyses (CNA), Professional Paper No. 27, 14 April 1970, 18 pp, UNCLASSIFIED, available from publisher and DTIC.

Fain, Janice B., "The Lanchester Equations and Historical Warfare: An Analysis of Sixty World War II Land Engagements," History, Numbers, and War, Vol 1, No 1 (Spring), 1977, pp 34-52. UNCLASSIFIED, available from libraries. Originally published as an unclassified paper contained in the Proceedings of the 34th Military Operations Research Symposium, December 3-5, 1974 (SECRET).

Helmbold, Robert L., "Historical Data and Lanchester's Theory of Combat," Combat Operations Research Group (CORG), HQ, US Continental Army Command, Staff Paper CORG-SP-128, 1 July 1961, 178 pp, UNCLASSIFIED, available from DTIC as AD-480-975.

Helmbold, Robert L., "A Modification of Lanchester's Equations," Operations Research, Vol 13, No 5, Sep-Oct 1965, pp 857-859. UNCLASSIFIED, available from libraries.

Helmbold, Robert L., "Some Observations on the Choice of Exponent in Lanchester's Law," unpublished manuscript, presented to a meeting of the Ad Hoc Working Group on the Theory of Combat, Cosponsored by the Defense Nuclear Agency and the Naval Postgraduate School, held at Monterey, CA, 1979.

Helmbold, Robert L., "Combat History Analysis Study Effort (CHASE): Progress Report for the Period August 1984-June 1985," US Concepts Analysis Agency, 8120 Woodmont Avenue, Bethesda, MD 20814-2797, Concepts Analysis Agency Technical Paper CAA-TP-86-2, August 1986, 249 pp, UNCLASSIFIED, available from DTIC as AD-F860-122.

Helmbold, Robert L., "Do Battles and Wars Have a Common Relationship Between Casualties and Victory?," US Army Concepts Analysis Agency, 8120 Woodmont Avenue, Bethesda, MD, 20814-2797, Concepts Analysis Agency Technical Paper CAA-TP-87-16, November 1987, 52 pp, UNCLASSIFIED, available from DTIC as AD-A196-126.

Helmbold, Robert L., "Rates of Advance in Historical Land Combat Operations," US Army Concepts Analysis Agency, 8120 Woodmont Avenue, Bethesda, MD 20814-2797, Concepts Analysis Agency Research Paper CAA-RP-90-1, June 1990, 142 pp, UNCLASSIFIED, available from DTIC as AD-A225-635.

Helmbold, Robert L., "Database of Battles: Version 1990 (Computer Diskette)," US Army Concepts Analysis Agency Data Base, 30 April 1991. AD-M000-121.

Helmbold, Robert L., Personnel Attrition Rates in Historical Land Combat Operations: A Catalog of Attrition and Casualty Data Bases on Diskettes Usable With Personal Computers, CAA Research Paper CAA-RP-93-4, September 1993, AD-A279-069. Also, PAR Data Disks, diskettes accompanying the preceding research paper, AD-M000-344 (compressed Quattro Pro format). Revised set of diskettes, AD-M000-368 (uncompressed Lotus 1-2-3 format).

Helmbold, Robert L.,and Rehm, Allan S., "The Influence of the Numerical Strength of Engaged Forces on Their Casualties," (a translation of Osipov's papers), *Naval Research Logistics Journal, Special Issue on Air-Land-Naval Warfare Models, Part II*, April 1995, Vol 42, pp 435-490, 1995.

Lanchester, Frederick William, *Aircraft in Warfare: The Dawn of the Fourth Arm*, Constable & Co., London, 1916. UNCLASSIFIED, available from libraries. The portion dealing with Lanchester's square law was reprinted in The World of Mathematics, James R. Newman (ed), Vol 4, pp 2138-2157, published by Simon & Schuster, New York, 1956.

Osipov, M., "The Influence of the Numerical Strength of Engaged Forces on Their Casualties," (In Russian), Military Collection, Numbers 6 through 10, 1915, (No 6, pp 59-74; No 7, pp 25-36; No 8, pp 31-40; No 9, pp 25-37; No 10, pp 93-96). UNCLASSIFIED, available from libraries. English translation published as US Army Concepts Analysis Research Paper CAA-RP-91-2, September 1991, and available from DTIC. English translation also published as Helmbold, Robert L. and Rehm, Allan S., "The Influence of the Numerical Strength of Engaged Forces on Their Casualties," *Naval Research Logistics Journal, Special Issue on Air-Land-Naval Warfare Models, Part II*, April 1995, vol 42, pp 435-490, 1995.

Peterson, Richard H., "On the 'Logarithmic Law' of Attrition and Its Application to Tank Combat," Operations Research, Vol 15, no 3, May-June 1967, pp 557-558. UNCLASSIFIED, available from libraries.

Small, Melvin; and Singer, J. David, *Resort to Arms: International and Civil Wars, 1816-1980*, Sage Publications, Beverly Hills, 1982, ISBN: 0-8039-1776-7, 1982, 372 pp, UNCLASSIFIED, available from publisher and libraries.

Taylor, James G., *Lanchester Models of Warfare*, in two volumes, Military Applications Section, Operations Research Society of America, March 1983, UNCLASSIFIED, available from libraries.

Willard, Daniel, "Lanchester as Force in History," Research Analysis Corporation (RAC), Technical Paper RAC-TP-74, November 1962, 37 pp, UNCLASSIFIED, available from DTIC as AD-297-375.

APPENDIX D

TECHNICAL NOTES

D-1. INTRODUCTION. Paragraphs D-2 through D-5 of this appendix review some basic concepts underlying the use of Gaussian kernel smoothers. As such, they provide some background and motivation for the use of Gaussian kernel smoothers in this study. Paragraph D-6 provides a guide to some of the most important sources of information on the origins and development of the advantage parameter concept and its validation against historical data.

D-2. STATISTICAL MODEL OF THE PROBABILITY OF WINNING

a. We begin by noting that, for most of this report, the data on a given battle or war consists of the following:

(1) An identification of which side was the attacker and which was the defender.

(2) The strengths and losses on both sides.

(3) An identification of which side won.

b. From these strengths and losses, we compute the advantage parameter and the force ratio for each battle or war.

c. Our aim in this report is to examine how the probability of winning varies with the values of either the advantage parameter or the force ratio. However, the probability of winning is not given directly in the data. Accordingly, a statistical model of how the probability of winning is related to the given data is required. To develop such a statistical model, begin by letting *QZ* be some quantity that might be expected to affect the probability that side Z wins. For example, *QZ* might be the advantage parameter favoring side Z or the force ratio favoring side Z. Let *WINZ* be the indicator of whether side Z won. That is, $WINZ = 1$ if side Z won, and $WINZ = 0$ if side Z lost. Note that *QZ* and *WINZ* are provided directly by the available data. The statistical model postulates that there is a functional relationship between the observed quantity *QZ* and the probability that side Z wins. Mathematically, this is expressed as

$$\Pr(WINZ) \equiv \Pr(WINZ = 1 | QZ) = F(QZ) \qquad \text{(D-1)}$$

where $\Pr(WINZ = 1 | QZ)$ is the probability that side Z wins when side Z has the advantage characterized by the quantity *QZ* and $F(QZ)$ is some function of *QZ* that gives the value of this probability. It follows mathematically that the probability that side Z loses is given by the complementary probability

$$\Pr(WINZ = 0 | QZ) = 1 - \Pr(WINZ) = 1 - F(QZ) \qquad \text{(D-2)}$$

D-3. METHODS FOR ASSESSING THE PROBABILITY OF WINNING

a. The statistical literature provides a number of methods for estimating the function $F(QZ)$. One class of these methods postulates that $F(QZ)$ has a specific functional form, usually involving a few parameters which are to be fitted to the data. These are called parametric statistical methods. However, in this report we want to proceed without any special assumptions regarding the functional form of $F(QZ)$. Accordingly, we use nonparametric statistical methods. (To simplify the presentation, all of the examples in this Appendix take $QZ = ADVY$ and $WINZ = WINY$. Accordingly, we write $F(ADVY) = \Pr(WINY = 1 | ADVY)$. However, similar methods and observations also apply when QZ is taken to be the force ratio or any other quantity that is fully determined by the given data and that might affect which side wins.)

b. Perhaps the easiest nonparametric method is simply to list the battles in order by the quantity *ADVY*, and to see whether there is a tendency for side Y to lose almost all the battles with small values of *ADVY*, to win a larger and larger fraction of battles as *ADVY* increases, and to win almost all the battles with high values of *ADVY*. Table D-1 illustrates this approach. It was constructed using the following procedure. First, we selected all of the battles in the CDB90DAT data base that had clearly-defined initial strengths and no reinforcements. This provided us with a sample of 290 battles. We computed the advantage parameter favoring the defender (*ADVY*) for each of these battles. We then listed the battles in order of increasing *ADVY* value, and recorded the corresponding observed victorious side (*WINY* = 0 or 1). Then, merely to keep the table reasonably short, we extracted every tenth battle. This resulted in the 29 battles listed in Table D-1. As can be seen, when the defender's advantage is low, the defender loses almost all the battles. When the defender's advantage is high, the defender wins nearly all the battles. And as the defender's advantage increases, the defender tends to win more and more frequently.

c. The simple listing method is easy to understand. As illustrated in Table D-1, it can quickly give a rough indication of whether the probability of winning really does depend on the independent variable, which in this case is *ADVY*. However, it has several limitations. When the number of battles is very large, the list becomes too long to justify presenting all of it. Also, the interpretation of such a list is often rather subjective, and in fact becomes very difficult when the probability of winning is only weakly related to the independent variable. A better approach to estimating the function *F*(*ADVY*) might be to convert the simple list into a contingency table form. For example, in Table D-1, we note that the defender lost 14 out of the 15 battles with an *ADVY* of –0.22 or less, won 7 out of the 7 battles with an *ADVY* of +0.22 or more, and won 3 of the 7 battles with an *ADVY* between –0.22 and +0.22. For all 290 battles, we can construct the contingency table shown in Table D-2. This table shows that the defender won only 112 (38.6 percent) of the 290 battles. However, the defender won only 2 (6.7 percent) of the 30 battles with *ADVY* less than –1.0. On the other hand, the defender won 13 (100.0 percent) of the 13 battles with *ADVY* over +1.0. And, by and large, the percentage of battles won by the defender increased as the defender's advantage parameter increased. We can apply the conventional statistical chi-square test for independence in contingency tables to these data. We get a chi-square value of 127.6 at 6 degrees of freedom, so that the probability of a defender victory is almost certainly dependent on the defender's advantage parameter—but that was already evident

from a visual inspection of Table D-2, so these chi-square results add little or nothing to our understanding of what is going on.

Table D-1. Simple List of *WINY* versus *ADVY*

Row no.	*ADVY*	*WINY*
1	-1.57	0
2	-1.29	0
3	-1.12	1
4	-0.88	0
5	-0.81	0
6	-0.72	0
7	-0.64	0
8	-0.59	0
9	-0.54	0
10	-0.51	0
11	-0.43	0
12	-0.37	0
13	-0.33	0
14	-0.28	0
15	-0.22	0
16	-0.13	1
17	-0.06	0
18	-0.01	1
19	0.01	0
20	0.05	0
21	0.07	1
22	0.17	0
23	0.24	1
24	0.33	1
25	0.41	1
26	0.48	1
27	0.65	1
28	0.90	1
29	1.16	1

Table D-2. Cross-Tabulation of *WINY* versus *ADVY*

ADVY bracket	*WINY* = 0	*WINY* = 1	Total	Percent won by defender (side Y)
Less than -1.0	28	2	30	6.7
-1.0 to -0.6	42	1	43	2.3
-0.6 to -0.3	49	6	55	10.9
-0.3 to +0.3	53	48	101	47.5
+0.3 to +0.6	5	25	30	83.3
+0.6 to +1.0	1	17	18	94.4
Over +1.0	0	13	13	100.0
Total	178	112	290	38.6

d. While the cross-tabulation presentation does provide more insight than a simple list, it has certain shortcomings. One of these is that slightly different results would be obtained if either different-sized brackets or a different number of brackets were used. A more serious objection is that the smaller the brackets are made, the fewer the battles they contain and hence the more variable is the percentage won by the defender. So, after the bracket size has been reduced to a certain point, the "signal" (that is, the relation between *ADVY* and Pr(*WINY*)) simply becomes lost in the noise. Another difficulty with the cross-tabulation presentation is that, when the brackets are made broad enough to get a high signal-to-noise ratio, they also become so broad as to obscure a lot of potentially important detail. For instance, it's hard to tell from Table D-2 just how the function $F(ADVY)$, describing the probability the defender wins, varies with *ADVY*. A simple moving average might be tried to overcome some of these problems. Figure D-1 shows how this works out for the 290-battle extract from the CDB90DAT data base. Here a 15-point moving average is used. Attempts to use more or fewer points led to less attractive results (fewer points led to a more "noisy" picture, while more seemed to over-smooth the data). This display clearly shows the strong rise in the probability the defender wins as the defender's advantage increases.

e. The moving average smoother has a number of advantages over the cross-tabulation method, and is easy to understand. However, it, too, has certain disadvantages. For example, the remaining "noise" or choppy variations and jumps in the moving average line makes it uncertain whether the probability increases smoothly and steadily, but certainly does not rule out that possibility. Also, because the moving average process unavoidably introduces statistical dependencies and correlations among the smoothed values, its statistical analysis requires rather delicate and sophisticated treatment. The conventional moving average essentially assigns the *same* weight to each point in its moving "window." A less "noisy" presentation can be obtained by assigning *unequal* weights to the points in the moving "window." One method of assigning unequal weights is to use a kernel smoother. If $K(s)$ is a kernel function, then the kernel smoother of the n observations (x_i, y_i) is defined to be the following weighted average of the data values:

$$\bar{y}(x) = \frac{\sum_{i=1}^{n} y_i K(x - x_i)}{\sum_{i=1}^{n} K(x - x_i)} \tag{D-2}$$

Here x varies over the x-values of interest, and $\bar{y}(x)$ is the smoothed value corresponding to the point x. Kernel functions are often required to be nonnegative, symmetric in s (that is, that $K(s) = K(-s)$), have a maximum at $s = 0$, and to approach zero as s tends to infinity. A Gaussian kernel, defined by

$$K(s) = \exp[-(s/\sigma)^2/2] \tag{D-3}$$

satisfies all these requirements. We refer to the parameter σ in equation (D-3) as the standard deviation of the Gaussian kernel. It plays a role analogous to the number of points used in a moving average smoother. The larger the standard deviation the more weight the Gaussian kernel places on remote points. Figure D-2 shows how this works out when the 290-battle extract from the CDB90DAT data base is smoothed with a Gaussian kernel using a standard deviation h equal to 0.10. Random variation or "noise" still causes some "wiggles" in the smoothed presentation, but it is considerably smoother and easier to interpret than the moving average presentation. Other values of the standard deviation were tried, but a value of about 0.10 seemed on balance to give the best results. Accordingly, we elected to smooth all the data using a Gaussian kernel with a standard deviation of 0.10.

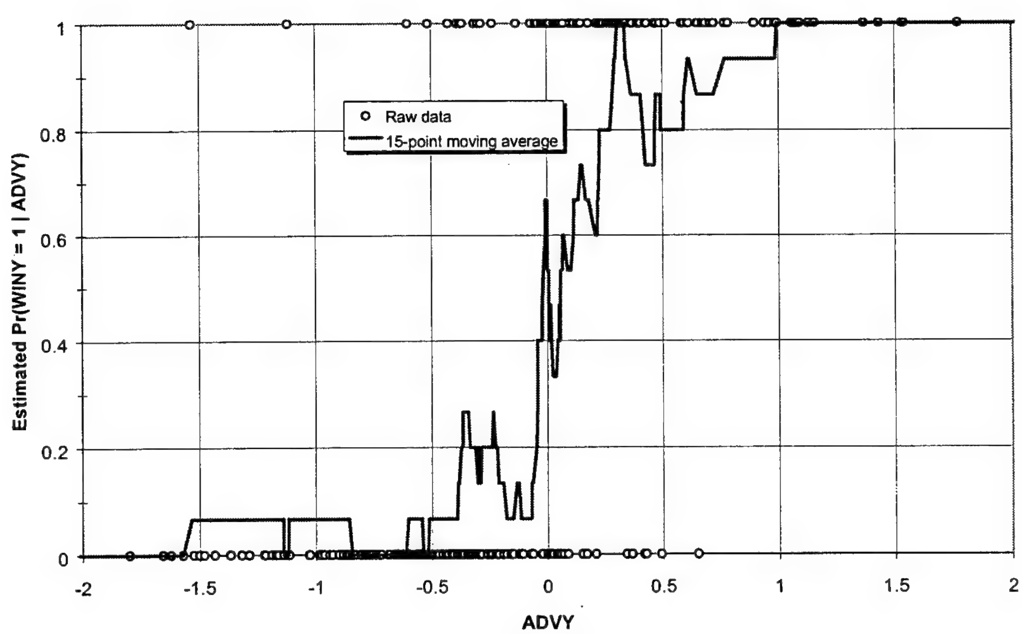

Figure D-1. Simple Moving Average of *WINY* versus *ADVY*

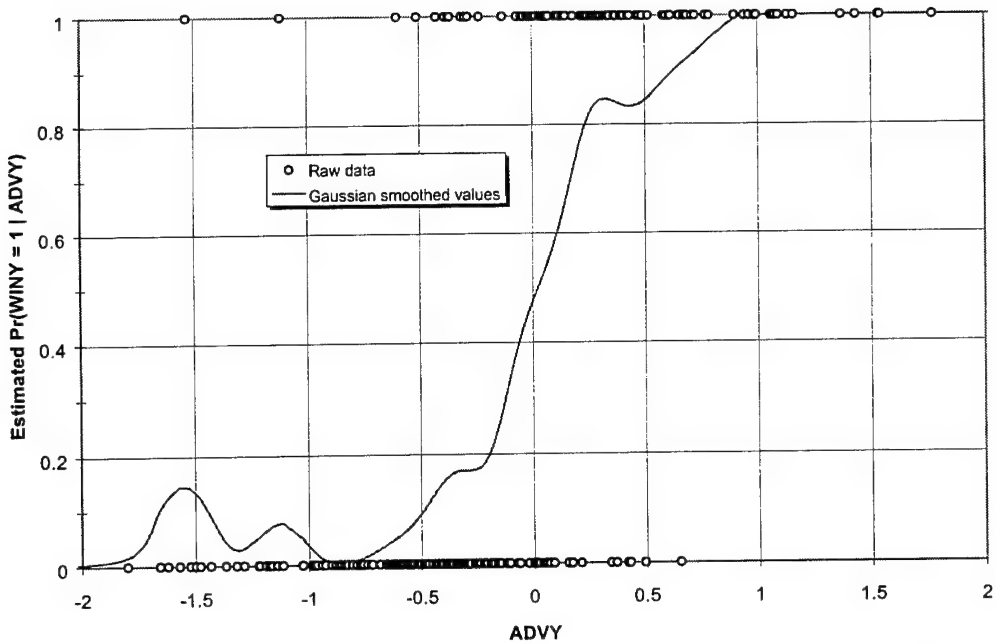

Figure D-2. Gaussian Kernel Smooth of *WINY* versus *ADVY*

f. It is true, of course, that kernel smoothers, including Gaussian smoothers, have certain disadvantages. They share with moving averages the characteristic of introducing statistical dependencies and correlations between the estimated or smoothed data points, and a sensitivity to the choice of standard deviation (or width of the smoothing window).

D-4. SOME EXPERIMENTS ON SMOOTHING KNOWN FUNCTIONS

a. In reality, we do not know the functional form of $F(QZ) = \Pr(WINZ = 1|QZ)$. However, it is still desirable to know how well a Gaussian kernel smoother can recover a known $F(QZ)$. For such experiments, we artificially dictate a particular functional form for $F(QZ) = \Pr(WINZ = 1|QZ)$. These probabilities are then used in a Monte Carlo simulation to generate a random sample of data. These simulated data are then smoothed using a Gaussian kernel smoother, and the results compared to the dictated form of $F(QZ)$. In all our experiments, we used the following dictated or known functional form for $F(QZ)$

$$F(q) = \frac{\exp(q_0 + kq)}{1 + \exp(q_0 + kq)} + h\exp\left(-(q - q_1)^2 / b\right) \tag{D-4}$$

The first term on the right-hand side of equation D-4 represents a logistic curve that crosses the 50-percent value at $q = -q_0 / k$. The second term on the right-hand side represents a Gaussian-shaped "bump" centered at $q = q_1$ and having a breadth b and height h. The parameters (q_0, k, h,

b, and q_1) in this formula are introduced to allow some flexibility in the specific functional form used for $F(QZ)$.

b. Figure D-3 shows the results of one experiment of this type. Here a sample of 300 data points was generated by Monte Carlo simulation from the probability function of equation D-4 with parameter values $q_0 = 0.00$, $k = 2.50$, $q_1 = -1.00$, $b = 0.05$, and $h = 0.10$. In Figure D-3, the dictated function is represented by the curve labeled "True Pr(*WINZ*)" and the simulated data points are labeled "Simulated *WINZ*." Both a simple moving average and a Gaussian kernel smoothing of the simulated data are shown. While the Gaussian smoother does much better than the simple moving average at recovering the dictated functional form, it is clear that it still is only a rough approximation to it.

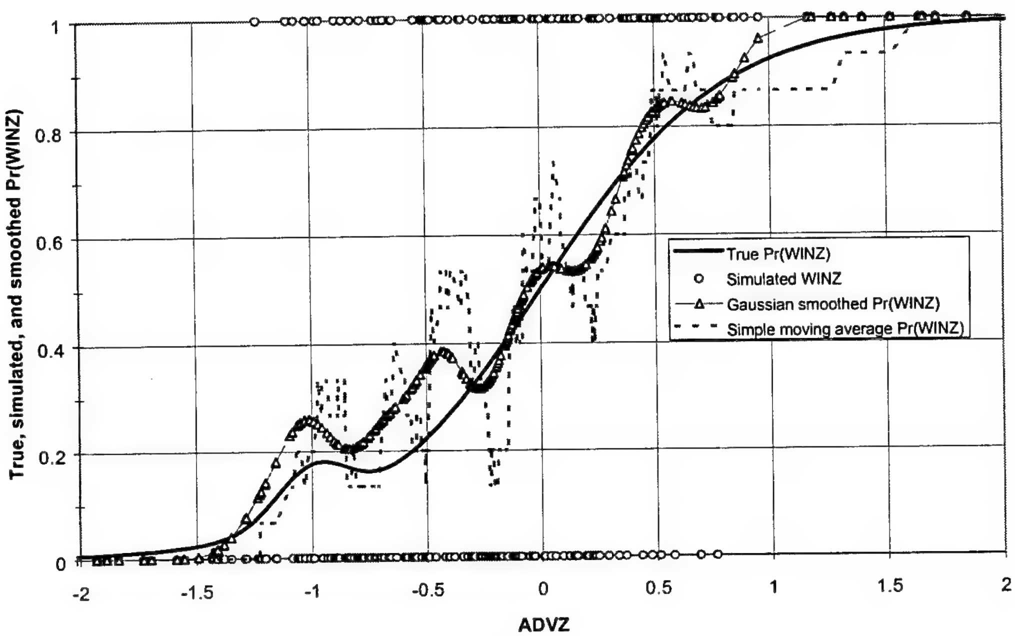

Figure D-3. A Simulation Experiment With Gaussian and Moving Average Smoothing

c. Figure D-4 shows the results of four separate simulation runs. In each of these runs, the dictated probability function is given by equation D-4 with parameter values $q_0 = 0.00$, $k = 2.50$, $q_1 = -1.00$, $b = 0.05$, and $h = 0.00$. Note that in this case the dictated probability function is just a logistic function with no "bump," because $h = 0.00$. Nevertheless, as indicated in Figure D-4, in three out of the four runs the Gaussian smooth has a bump near $ADVZ = -1.5$, and in two of the four runs has a bump near $ADVZ = -1$. Accordingly, whether the bumps observed in the actual data presented in the body of this report and also illustrated in Figure D-2 are "real" or merely an artifact of the smoothing process is at present uncertain.

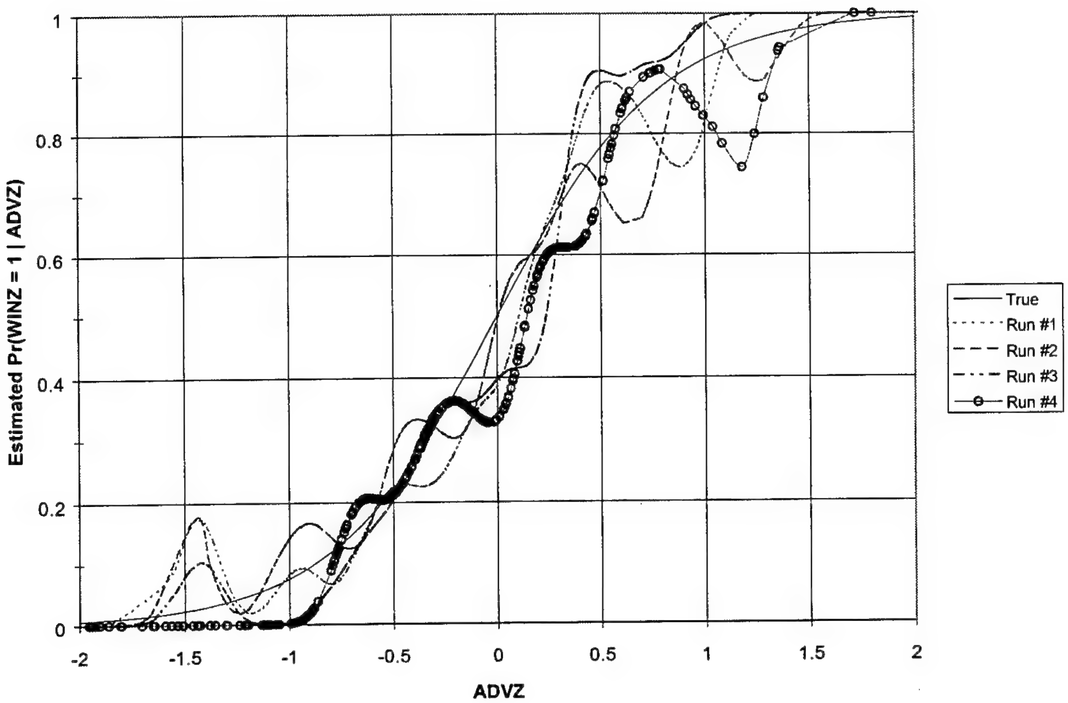

Figure D-4. Gaussian Smoothing of Four Separate Simulated Cases

D-5. SELECTED REFERENCES ON KERNEL SMOOTHING. Those who wish to study kernel smoothing in more detail may find some of the following references, or the works they cite, a satisfactory place to start.

a. Eubank, R., *Spline Smoothing and Nonparametric Regression*, Dekker, New York, 1988.

b. Fan, J., and Gijbels, I., *Local Polynomial Modelling and Its Applications*, Chapman and Hall, London, 1996.

c. Fan, J., Design-Adaptive Nonparametric Regression, *J. Amer. Statist. Assoc.*, **87**(1992), pp 998-1004.

d. Fan, J., Heckman, N. E., and Wand, M. P., Local Polynomial Kernel Regression for Generalized Linear Models and Quasi-Likelihood Functions, *J. Amer. Statist. Assoc.*, **90**(1995), no 429, pp 141-150.

e. Gasser, Theo , Kneip, Alois, and Köhler, Walter, A Flexible and Fast Method for Automatic Smoothing, *J. Amer. Statist. Assoc.*, **86**(1991), no 415 (Sep), pp 643-652.

f. Härdle, W., Hall, P., and Marron, J. S., Regression Smoothing Parameters That Are Not Far From Their Optimum, *J. Amer. Statist. Assoc.*, **87**(1992), no 417 (Mar), pp 227-233.

g. Hjort, N. L, and Jones, M. C., Locally Parametric Nonparametric Density Estimation, *Ann. Statist.*, **24**(1996), no 4(August), pp 1619-1647.

h. Jennen-Steinmetz, Ch. and Gasser, Th., A Unifying Approach to Nonparametric Regression Estimation, *J. Amer. Statist. Assoc.*, **83**(1988), pp 1084-1089.

i. Jones, M. C., Linton, O., and Nielsen, J. P., A Simple and Effective Bias Reduction Method for Density and Regression Estimation, *Biometrika*, **82**(1995), pp 327-338.

j. Jones, M.C., Marron, J.S., and Sheather, S. J., A Brief Survey of Bandwidth Selection for Density Estimation, *J. Amer. Statist. Assoc.*, **91**(1996), pp 401-407.

k. Kychka, Douglas, Choosing a Range for the Amount of Smoothing in Nonparametric Regression, *J. Amer. Statist. Assoc.*, **86**(1991), no 415 (Sep), pp 653-664.

l. Lall, U., Recent Advances in Nonparametric Function Estimation: Hydrologic Applications, US National Report to IUGG, 1991-1994, *Rev. Geophys*, **133** Supp 1.

m. Müller, H-G., Weighted Local Regression and Kernel Methods for Nonparametric Curve Fitting, *J. Amer. Statist. Assoc.*, 82 (1987), no 397 (Mar), pp 231-238.

n. O'Sullivan, F. , Yandell, B. , and Raynor, W. J., Automatic Smoothing of Regression Functions in Generalized Linear Models, *J. Amer. Statist. Assoc.*, **81**(1986), pp 96-104.

o. Schuchany, W. R., Adaptive Bandwidth Choice for Kernel Regression, *J. Amer. Statist. Assoc.*, **90**(1995), no 430, p535-540.

p. Staniswalis, J., The Kernel Estimate of a Regression Function in Likelihood-Based Models, *J. Amer. Statist. Assoc.*, **84**(1989), 276-283.

q. Staniswalis, J., Local Bandwidth Selection for Kernel Estimates, *J. Amer. Statist. Assoc.*, **84**(1989), no 405, pp 284-288.

r. Stone, C. J., Consistent Nonparametric Regression, *Ann. Statist.*, **5**(1977), pp 595-620.

s. Tibshirani, R., and Hastie, T., Local Likelihood Estimation, *J. Amer. Statist. Assoc.*, **82**(1987), pp 559-567.

t. Wand, M. P., and Jones, M. C., *Kernel Smoothing*, Chapman and Hall, London, 1995.

D-6. GUIDE TO BACKGROUND MATERIAL ON THE ADVANTAGE PARAMETER . This paragraph provides a guide to some of the most important sources of information on the origins and development of the advantage parameter concept and its validation against historical data.

a. The first work to derive the advantage parameter from Lanchester's equations and to study its empirical validation was Helmbold, Robert L., "Historical Data and Lanchester's Theory of Combat," Combat Operations Research Group (CORG), HQ, US Continental Army Command, Staff Paper CORG-SP-128, 1 July 1961, 178 pp, UNCLASSIFIED, available from DTIC as AD-480-975.

b. The next work, done to confirm the results of the first by exploring a second and independent data base of battles, was Helmbold, Robert L., "Historical Data and Lanchester's Theory of Combat: Part II," Combat Operations Research Group (CORG), HQ, US Combat Development Command, Staff Paper CORG-SP-190, August 1964, 132 pp, UNCLASSIFIED, available from DTIC as AD-480-109.

c. One of the results documented in the foregoing two works was published in Helmbold, Robert L., "Some Observations on the Use of Lanchester's Theory for Prediction," *Journal of the Operations Research Society of America*, Vol 12, No 5, 1964, pp 778-781. UNCLASSIFIED.

d. Some thoughts on the form of Lanchester's equations, inspired by the findings in the foregoing works, appeared as Helmbold, Robert L., "A Modification of Lanchester's Equations," *Operations Research*, Vol 13, no 5, Sep-Oct 1965, pp 857-859. UNCLASSIFIED.

e. An application of the results of the foregoing to validating combat models can be found in Helmbold, Robert L., "Some Observations on Validating Combat Models," Presented to the NATO Conference on Recent Developments in Lanchester Theory, Munich, Germany, 3-7 July 1967, 1967, 16 pp, UNCLASSIFIED.

f. A paper on the weakness of force ratio as a gauge of victory in battle is contained in Helmbold, Robert L., "Probability of Victory in Land Combat as Related to Force Ratio," The RAND Corporation, RAND Paper P-4199, October 1969, 14 pp, UNCLASSIFIED, available from DTIC as AD-696-489.

g. An analysis of air battle data, using the same methods as in items D-6a and D-6b, can be found in Helmbold, Robert L., "Air Battles and Ground Battles–A Common Pattern?," The RAND Corporation, RAND P-4548, January 1971, 15 pp, UNCLASSIFIED, available from DTIC as AD-718-975.

h. A critique of the conventional breakpoint methods for terminating simulated battles and engagements is given in Helmbold, Robert L., "Decision in Battle: Breakpoint Hypotheses and Engagement Termination Data," The RAND Corporation, RAND Report R-772-PR, June 1971, 96 pp, UNCLASSIFIED, available from publisher and DTIC as AD-729-769).

i. An application of the advantage parameter to modeling theater escalation and de-escalation decisions appeared in "Integrated Warfare—Representing the Decision Processes Within Computer Simulations of Combat," US Army Concepts Analysis Technical Paper CAA-TP-85-1, January 1985, UNCLASSIFIED, available from DTIC.

j. A validation of the predictions made in the CORG papers (see paragraphs D-6a and D-6b), using a much larger and more extensive data base of battles than had previously been available, is given in Helmbold, Robert L., "Combat History Analysis Study Effort (CHASE): Progress Report for the Period August 1984-June 1985," US Concepts Analysis Agency, 8120 Woodmont Avenue, Bethesda, MD 20814-2797, Concepts Analysis Agency Technical Paper CAA-TP-86-2, August 1986, 249 pp, UNCLASSIFIED, available from DTIC as AD-F860-122. This work also introduced the technique of logistic regression to analyze the data, which was a considerable advance over the less sophisticated methods used in earlier works.

l. A test of the extrapolatability to wars of the predictions based on battle data in works such as those of paragraphs D-6a, D-6b, D-6f, and D-6i is contained in Helmbold, Robert L., "Do Battles and Wars Have a Common Relationship Between Casualties and Victory?," US Army Concepts Analysis Agency, 8120 Woodmont Avenue, Bethesda, MD, 20814-2797, Concepts Analysis Agency Technical Paper CAA-TP-87-16, November 1987, 52 pp, UNCLASSIFIED, available from DTIC as AD-A196-126.

m. An application of the advantage parameter to analyzing rates of advance is given in Helmbold, Robert L., "Rates of Advance in Historical Land Combat Operations," US Army Concepts Analysis Agency, 8120 Woodmont Avenue, Bethesda, MD 20814-2797, Concepts Analysis Agency Research Paper CAA-RP-90-1, June 1990, 142 pp, UNCLASSIFIED, available from DTIC as AD-A225-635. Much of this work also appears in Helmbold, Robert L., "Rates of Advance in Land Combat Operations," *Naval Research Logistics Journal, Special Issue on Air-Land-Naval Warfare Models, Part III*, Volume 42, No 4, June 1995, pp 635-670.

n. The data bases of battles analyzed in this report are described in the following works. The CDB90DAT data base is described in Helmbold, Robert L., "Database of Battles: Version 1990 (Computer Diskette)," US Army Concepts Analysis Agency Data Base, 30 April 1991. AD-M000-121. The other data bases are described in Helmbold, Robert L., Personnel Attrition Rates in Historical Land Combat Operations: A Catalog of Attrition and Casualty Data Bases on Diskettes Usable With Personal Computers, CAA Research Paper CAA-RP-93-4, September 1993, AD-A279-069. Also, PAR Data Disks, diskettes accompanying the preceding research paper, AD-M000-344 (compressed Quattro Pro format). Revised set of diskettes, AD-M000-368 (uncompressed Lotus 1-2-3 format).

APPENDIX E

DISTRIBUTION

Addressee	No of copies
Deputy Under Secretary of the Army (Operations Research) ATTN: ODUSA-OR 102 Army Pentagon Washington, DC 20310-0102	1
Deputy Chief of Staff for Operations and Plans Headquarters, Department of the Army ATTN: DAMO-ZXA 400 Army Pentagon Washington, DC 20310-0400	1
Deputy Chief of Staff for Logistics Headquarters, Department of the Army ATTN: DALO-ZA Room 3E560, The Pentagon Washington, DC 20310-0500	1
Deputy Chief of Staff for Personnel Headquarters, Department of the Army ATTN: DAPE-ZX 300 Army Pentagon Washington, DC 20310-0300	1
Director US Army TRADOC Analysis Command-White Sands ATTN: ATRC-W Bldg 1400 1400 Martin Luther King Drive White Sands Missile Range, NM 88002-5502	1
TRAC-OAC ATTN: ATRC-FSV 255 Sedgewick Ave Fort Leavenworth, KS 66027-2345	1

Addressee	No of copies
HQ TRADOC Deputy Chief of Staff for Simulations and Analysis ATTN: ATAN-S Bldg. 5G, Room 303G Fort Monroe, VA 23651-5143	1
Director US Army Materiel Systems Analysis Activity ATTN: AMXSY-D 392 Hopkins Road Aberdeen Proving Ground, MD 21005-5071	1
Defense Technical Information Center ATTN: BCP Product Management Branch 8725 John J. Kingman Road, STE 0944 Ft. Belvoir, VA 22060-6218	2
USASCAF The Pentagon Library ATTN: JDHQ-LR (Army Studies) 6605 Army Pentagon Washington, DC 20310-6605	1
Director Program Analysis & Evaluation (PA&E) Office of the Secretary of Defense Room 2E330, The Pentagon Washington, DC 20310-1800	1
Integration and Assessment Division Joint Staff/J8 (COL R. Jones) Room 1D964, The Pentagon Washington, DC 20318-8000	1
Joint Chiefs of Staff SJCS, IMD ATTN: R&A 400 Joint Staff Pentagon Washington, DC 20318-0400	1
Commandant US Army War College ATTN: AWCSL (Library) 122 Forbes Road Carlisle Barracks, PA 17013-5050	1

Addressee	No of copies
US Army War College Strategic Studies Institute ATTN: Librarian 122 Forbes Road Carlisle Barracks, PA 17013-5050	1
Air University ATTN: AUL/LD 600 Chennault Circle Maxwell Air Force Base, AL 36112-6424	1
Air War College ATTN: CADRE/WGO 401 Chennault Circle Maxwell Air Force Base, AL 36112-6428	1
US Navy War College ATTN: Code 1E11 (Library) 686 Cushing Road Newport, RI 02841-1207	1
President National Defense University ATTN: NDU-LD (Library) Bldg 62, 300 5th Avenue Ft McNair, DC 20319-5066	1
Commandant Armed Forces Staff College ATTN: Library/62 7800 Hampton Blvd Norfolk, VA 23511-1702	1
US Army Command and General Staff College Combined Arms Research Library 250 Gibbons Avenue Fort Leavenworth, KS 66027-2314	1
United States Military Academy DOIM Publications ATTN: MAIM-DS-RT Bldg. 667A, 2nd Floor Ruger Road West Point, NY 10996-1583	1

Addressee	No of copies
Naval Postgraduate School Dudley Knox Library ATTN: Greta Marlatt 411 Dyer Road Monterey, CA 93943-5000	1
Commander US Army, Pacific ATTN: APOP-PL Fort Shafter, HI 96858-5100	1
Combined Forces Command C-3 Operations Branch PSC 303, Box 27 FKJ3-PL-OA APO AP 96204-0027	1
US Army National Ground Intelligence Center ATTN: IANG-ML 220 Seventh St. NE Charlottesville, VA 22901-5396	1
Director US Army Engineer Studies Center Casey Building, No. 2594 ATTN: CETEC-ES 7701 Telegraph Road Alexandria, VA 22315-3803	1
Commander in Chief US Army, Europe ATTN: AEAGX-OR Unit 29351 APO AE 09014	1
Headquarters Air Combat Command Studies and Analysis Squadron 204 Dodd Blvd Suite 202 Langley Air Force Base, VA 23665-2778	1
Air Force Institute of Technology Library (LDC) Bldg 642, Area B 2950 P St Wright-Patterson AFB, OH 45433	1

Addressee	No of copies
Chief of Naval Operations ATTN: N-09B34F2 Bldg 36 901 M St SE Washington Navy Yard, DC 20374	1
Office of the Chief of Naval Research ATTN: Code 09M Arlington, VA 22217-5660	1
US Army Center of Military History ATTN: DAMH-RAM 1099 14th St. NW Washington, DC 20005-3402	1
Air Force Studies and Analyses Agency AFSAA/SAGC 1570 Air Force Pentagon Washington, DC 20330-1570	1
SHAPE Technical Centre Operations Research Division The Hague, Netherlands APO New York 09159	1
The RAND Corp. ATTN: R. Darilek 2100 M St. NW Washington.DC 20037-1270	1
Director RAND Arroyo Center 1700 Main Street Santa Monica, CA 90407	1
Center for Naval Analyses ATTN: Document Control 4401 Ford Avenue Alexandria, VA 22302-1498	1
Institute for Defense Analysis ATTN: TISO 1901 North Beauregard St. Alexandria, VA 22311-1772	1

Addressee	No of copies
MITRE Corp ATTN: Dr. R. Richards (Mail Stop 558) 1820 Dolly Madison Blvd McLean, VA 22102-3481	1
Internal Distribution:	
Unclassified Library	2
Originating office (CSCA-TA)	2
Author	20

GLOSSARY

Mathematical Symbols:

x = instantaneous personnel strength of the attacker

y = instantaneous personnel strength of the defender

$a = x / x_0 =$ instantaneous surviving fraction of side X (the attacker)

$d = y / y_0 =$ instantaneous surviving fraction of side Y (the defender)

A = Lanchester attrition coefficient, in $dy / dt = -Ax$

D = Lanchester attrition coefficient, in $dx / dt = -Dy$

FRX = initial force ratio favoring side X (the attacker) = x_0 / y_0

FRY = initial force ratio favoring side Y (the defender) = $y_0 / x_0 = 1 / FRX$

$\alpha = Ax_0 / y_0 = A \div FRY = A \times FRX$ = dimensionless attrition coefficient (as in equation (2-3))

$\delta = Dy_0 / x_0 = D \times FRY = D \div FRX$ = dimensionless attrition coefficient (as in equation (2-3))

$\mu^2 = \delta / \alpha = \dfrac{1-a^2}{1-d^2} =$ the mu-parameter arising from the solution given by equation (2-5) to the "dimensionless" Lanchester's square law of equation (2-2)

$\lambda = \sqrt{\alpha\delta} = \sqrt{AD} =$ the lambda-parameter arising from the solution given by equation (2-5) to the "dimensionless" Lanchester's square law of equation (2-2)

$f_x = 1 - a$ = instantaneous personnel casualty fraction to side X (the attacker)

$f_y = 1 - d$ = instantaneous personnel casualty fraction to side Y (the defender)

$ADVY = \ln(\mu) =$ the advantage parameter favoring side Y (the defender)

$WINY$ = index of victory favoring side Y (the defender), i.e., $WINY = 1$ if side Y (the defender) wins, otherwise $WINY = 0$

$C_x = x_0 - x =$ instantaneous casualties to side X (the attacker)

$C_y = y_0 - y =$ instantaneous casualties to side Y (the defender)

$CERY = C_x / C_y$ = instantaneous casualty exchange ratio favoring side Y (the defender)

$FERY = f_x / f_y = CERY \times FRY =$ the fractional exchange ratio favoring side Y (the defender)

T = time at which the battle ends